AAAS

Science



SPECIAL ISSUE

MANDALAY QUAKE

30 OCTOBER 2025

Can AI predict the workings of a complete cell?
p. 436

Chimpanzees can weigh the evidence
pp. 444 & 521

Neuronal rewiring wins Eppendorf & *Science* Prize
p. 482



Your bridge to successful innovation.

NEB's Customized Solutions Team is here to help, and serves as a bridge to the support and resources you need to ensure your success.

Creating the right partnership is essential when pioneering a new life science product. Every aspect of development – technical expertise, reagent optimization, manufacturing scale, turnaround time, reagent quality, and comprehensive logistical support – is vital for achieving your objectives. And in the regulated markets landscape, these challenges magnify, demanding an even more specialized approach.

Your Bridge to Successful Innovation

- Leverage NEB's 50 years of experience in enzymology and reagent manufacturing
- As an extension of your team, we prioritize a deep understanding of your objectives, work with you on an optimal solution, and help to anticipate your future needs
- Benefit from our ISO 9001- and 13485-certified processes and commitment to quality, as well as our GMP-grade* production facility, and specialized lyophilization facility for the highest quality production standards
- Access unparalleled support from our dedicated account managers, program managers, technical scientists and production teams
- We work closely with you on inventory management and global distribution through our network of NEB-owned subsidiaries, to ensure successful commercialization

NEB's Customized Solutions Team will help you access novel products, meet quality specifications, speed time to market, and streamline your supply chain, allowing you to focus more on what matters most — innovation.

Ready to start the discussion? Learn more at
www.neb.com/customizedsolutions.

** "GMP-grade" is a branding term NEB uses to describe products manufactured or finished at NEB's Rowley facility. The Rowley facility was designed to manufacture products under more rigorous infrastructure and process controls to achieve more stringent product specifications and customer requirements. Products manufactured at NEB's Rowley facility are manufactured in compliance with ISO 9001 and ISO 13485 quality management system standards. However, at this time, NEB does not manufacture or sell products known as Active Pharmaceutical Ingredients (APIs), nor does NEB manufacture its products in compliance with all of the Current Good Manufacturing Practice regulations.*

NEW ENGLAND Biolabs®

CONTENTS

30 OCTOBER 2025
VOLUME 390
ISSUE 6772



ON THE COVER



A Buddha statue lies face down in rubble in this photo taken a week after the 28 March earthquake in Mandalay, Myanmar. The earthquake caused fatalities across a broad swath of the country and beyond. This issue features papers exploring the risk and reasons behind the anomalously fast, sustained rupture. See the special section beginning on page 454. Photo: Reuters

A March 2025 earthquake caused severe damage at its epicenter in Mandalay, Myanmar, including to pagodas and other religious sites.

SPECIAL ISSUE

2025 Mandalay earthquake

INTRODUCTION

454 2025 Mandalay earthquake
—Angela Hessler and Sumin Jin

PERSPECTIVE

456 Mandalay earthquake pushes rupture limits
Fast rupture of Earth's crust breaks a long fault, expanding the damaged area —K. E. Bradley and J. A. Hubbard

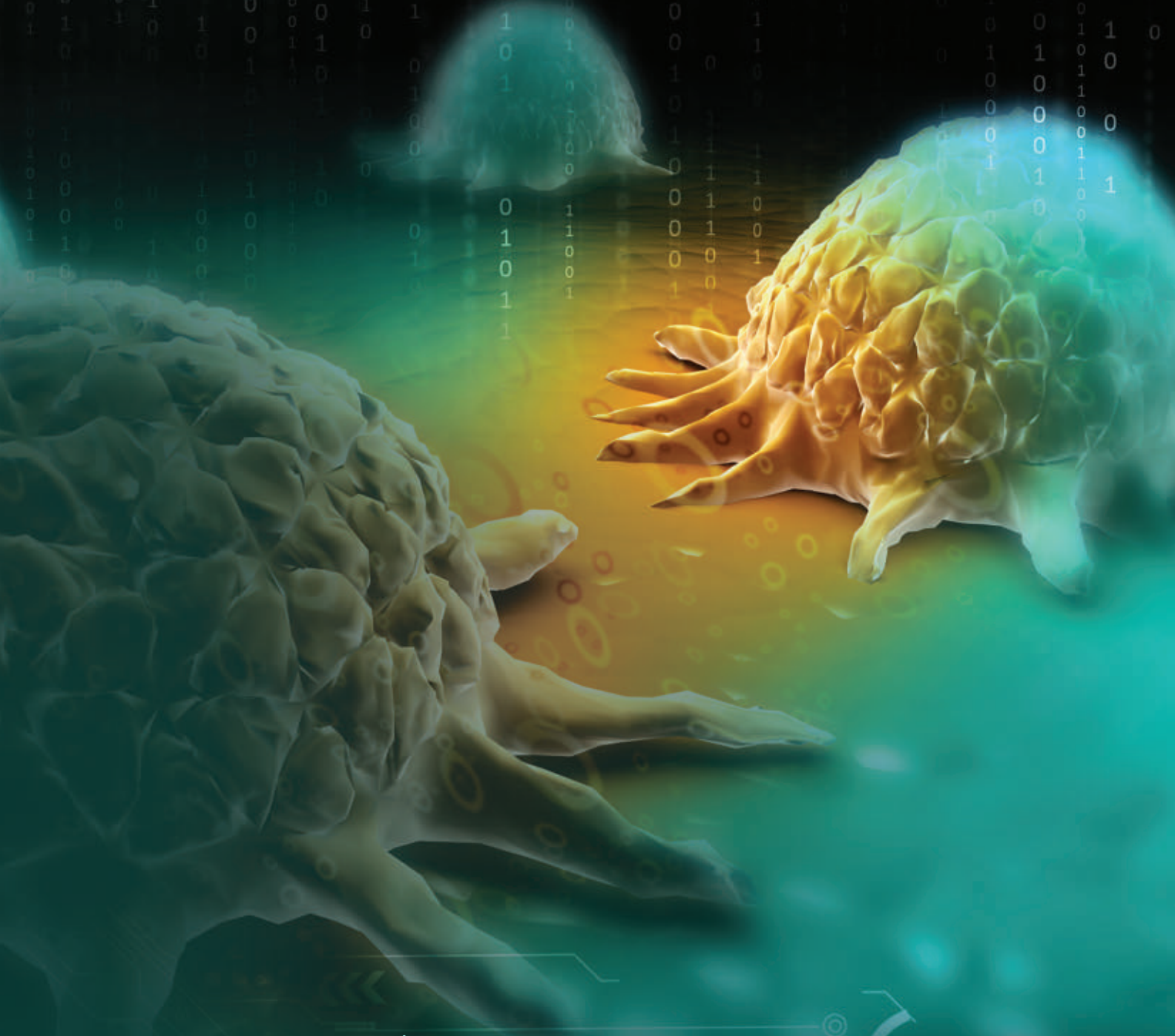
RESEARCH ARTICLES

458 Ultralong, supershear rupture of the 2025 M_w 7.7 Mandalay earthquake reveals unaccounted risk
—D. E. Goldberg *et al.*

463 Direct estimation of earthquake source properties from a single CCTV camera
—S. Latour *et al.*

468 Supershear rupture sustained through a thick fault zone in the 2025 M_w 7.8 Mandalay earthquake
—S. Wei *et al.*

476 Bimaterial effect and favorable energy ratio enabled supershear rupture in the 2025 Mandalay earthquake
—L. Xu *et al.*



2026

AAAS MARTIN AND ROSE WACHTEL CANCER RESEARCH AWARD

Recognizing the work of an early career scientist who has performed outstanding research in the field of cancer. Award nominees must have received their Ph.D. or M.D. within the last 10 years. The winner will deliver a public lecture on their research, receive a cash award of **$25,000**, and publish a Focus article on their award-winning research in *Science Translational Medicine*.

For more information visit
www.aaas.org/aboutaaas/awards/wachtel
or e-mail wachtelprize@aaas.org.
Deadline for submission: **February 1, 2026**.

AAAS Science Translational Medicine

EDITORIAL

425 Saving science by the sea
—S. Fitzpatrick and J. Maienschein

NEWS

426 Argentine telescope caught in U.S.-China crossfire
Superpower tensions have abruptly derailed a nearly finished radio telescope —M. de los Ángeles Orfila

427 Europe faults Big Tech platforms for lack of research access
Scientists welcome preliminary finding that Meta and TikTok are in breach of the Digital Services Act —K. Kupferschmidt

428 New antivenom is a 'potential game changer' for snakebites
Synthetic cocktail of llama and alpaca "nanobodies" protected mice —C. Wilcox

430 How NSF floats Antarctic research with no icebreaker
Ending *Palmer* lease is one of many belt-tightening moves amid budget uncertainty —J. Mervis

432 Dangerous ocean acidification tracked using ambient noise
Acoustic technique could make it easier to monitor threat to marine life stemming from rising carbon emissions —P. Voosen
PODCAST

434 At a novel meeting, AI wrote and reviewed the papers
Organizers aim to tune algorithms to reduce peer-review burdens and accelerate science —J. Brainard

FEATURES

436 The silicon cell
AI cell models could transform biomedicine—if they work as promised —M. Leslie

COMMENTARY

PERSPECTIVES

440 How common cuckoos adapt to multiple hosts
Genomic data reveal the complexity of egg mimicry evolution in cuckoos —M. D. Sorenson and C. N. Spottiswoode
RESEARCH ARTICLE p. 527

442 Broad benefits of the COVID-19 pandemic response
Nonpharmaceutical interventions protected the US population from other infectious diseases —D. N. Durrheim
RESEARCH ARTICLE p. 510

443 Two diverging paths for clean fuel
Fischer-Tropsch synthesis can produce liquid fuels from nonpetroleum sources —M. Saeys
RESEARCH SUMMARY p. 494; RESEARCH ARTICLE p. 516

444 Chimpanzees are natural scientists
Humans and chimpanzees share the potential to rationally revise their beliefs —B. Hare
RESEARCH ARTICLE p. 521

POLICY FORUM

446 How collaboration with China can revitalize US automotive innovation
Strategic collaboration, not protectionist barriers, should be pursued —J. P. Helveston

BOOKS ET AL.

449 More than a miracle
Pandemic preparedness is as essential as ever, argues a journalist —C. Rivers

450 Extraterrestrials illustrated
A physicist and a cartoonist join forces with an accessible guide to thinking about life in the cosmos —M. von Hippel

LETTERS

451 Asian dust threatens air pollution control efforts
—G. Wang *et al.*

452 Restore strict protection for wolves in Europe
—E. Randi *et al.*

452 Life in Science: A legendary request for narwhal DNA —E. Garde

PRIZE ESSAY

482 Respecifying partners
Neural circuits are rewired by manipulating cell surface combinatorial code —C. Lyu

REVIEWS

REVIEW SUMMARY

484 Materials science
Porous materials: The next frontier in energy technologies —E. M. Farber *et al.*

RESEARCH

HIGHLIGHTS

485 From *Science* and other journals

RESEARCH SUMMARIES

488 Solar cells
Electron accumulation across the perovskite layer enhances tandem solar cells with textured silicon —O. Er-raji *et al.*

489 Molecular biology
FIGNL1 inhibits homologous recombination in BRCA2 deficient cells by dissociating RAD51 filaments —R. Kuthethur *et al.*

490 Protein chaperones
Structural dissection of αβ-tubulin heterodimer assembly and disassembly by human tubulin–specific chaperones —Y. Seong *et al.*

491 Greenhouse gases
Global methane action pays for itself at least six times over —T. Stoerk *et al.*

492 Neuroimmunology
Regional encoding of enteric nervous system responses to microbiota and type 2 inflammation —P. Tan *et al.*

493 Neuroscience
Diverse somatic genomic alterations in single neurons in chronic traumatic encephalopathy —G. Dong *et al.*

494 Catalysis
Conversion of syngas into olefins with high hydrogen atom economy —C. Gao *et al.*
PERSPECTIVE p. 443; RESEARCH ARTICLE p. 516

RESEARCH ARTICLES

495 Molecular biology
Human RPA is an essential telomerase processivity factor for maintaining telomeres —S. Agrawal *et al.*

503 Polymer recycling
Glycolysis-compatible urethanases for polyurethane recycling —Y. Chen *et al.*

510 Epidemiology
Collateral effects of COVID-19 pandemic control on the US infectious disease landscape —T. S. Brett and P. Rohani
PERSPECTIVE p. 442

516 Catalysis
Trace-level halogen blocks CO_2 emission in Fischer-Tropsch synthesis for olefins production —Y. Cai *et al.*
PERSPECTIVE p. 443; RESEARCH SUMMARY p. 494

521 Comparative cognition
Chimpanzees rationally revise their beliefs —H. Schleihauf *et al.*
PERSPECTIVE p. 444

527 Pigmentation
Genomic architecture of egg mimicry and its consequences for speciation in parasitic cuckoos —J. Merondun *et al.*
PERSPECTIVE p. 440

WORKING LIFE

534 Adjusting to retirement
—T. F. Stuessy

424 *Science* Staff
533 *Science* Careers

Science **serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority or conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news, commentary, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.** *Science* (ISSN 0036-8075) is published weekly on Thursday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2025 by the American Association for the Advancement of Science. The title *Science* is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2865; Foreign postage extra: Air assist delivery: $135. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A. Change of address:** Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.science.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.



A high-mobility quadruped robotic dog, developed at KFUPM.

Nurturing entrepreneurship to drive economic success

Saudi Arabia's King Fahd University of Petroleum and Minerals (KFUPM) aims to create new economic sectors by attracting top talent and nurturing their entrepreneurial spirit.

KFUPM is in the midst of a complete transformation. "Not a single stone has been left unturned," says Muhammad M. Al-Saggaf, president of the university.

An alum of KFUPM and experienced industry leader, Al-Saggaf understands both where the university has been and where it needs to go. About a decade ago, as Saudi Arabia launched its economic development plan, Vision 2030, the government recognized that KFUPM would need to be an independent institution in order to reach its full potential. In 2020, Al-Saggaf was appointed as KFUPM president and began leading the university's transformation following its newfound independence. Since then, he has implemented sweeping changes to give KFUPM the flexibility and agility needed to strengthen both the local and global economies.

Diversity is a key component of the changes. That means diversity in thought, with more than 100 new interdisciplinary academic programs and more than 20 new interdisciplinary research centers. It also means embracing a rich cultural diversity with students coming from more than 75 countries. And although KFUPM was historically a male-only university for 60 years, the university also recognized the importance of gender diversity by opening the university's doors to women for the first time in 2021. At KFUPM, women now make up a remarkable 50% of undergraduate engineering students, well above the averages at many top institutions worldwide.

Science Custom Publishing sat down with Al-Saggaf to learn more about the evolution of KFUPM to a leading university that is transforming the research landscape and training the next generation of entrepreneurs to innovate in service of humanity.

KFUPM by the numbers



- ~100 new academic programs
- More than 20 new interdisciplinary research centers
- Students represent 75+ countries
- Ranks 67th globally overall
- Ranks 5th globally in # of patents
- 39% pursue graduate degrees
- Employability rate is 93%

PHOTO: PHOTOS PROVIDED BY KFUPM



Muhammad Al-Saggaf, President of KFUPM.

When you became president, you began a deliberate transformation of the university. What is KFUPM's current mandate?

We have a very definite mandate to help expand the economy of Saudi Arabia and, by implication, the global economy. But how do you expand the economy of a country that relies substantially on one sector, which is oil and gas? The answer is, you can only expand the economy by spawning new sectors, creating and nourishing them. To do this, you need two things: investment and talent, and one cannot make up for the other.

Our niche at KFUPM is developing economy-creating talent, the people who are going to create the future sectors, and subsequently the future jobs, and therefore contribute to the expansion of the economy. We tell our students when they join KFUPM, "You are not here to learn to get a job; you are here to learn to create a job." We know that this is achievable, as our students are some of the smartest students worldwide. The requirements to get into KFUPM are higher than those of medical schools, and we attract the top 0.8% of the talent in the region. With this, we have a stronghold on attracting the talent that are suitable to be economy-creating talent, to create the sectors of tomorrow.

How has research changed during the transformation?

We overhauled research 100% at KFUPM. The objective of research is not to produce publications and glory for the researcher or the institution, but rather to help humanity, to impact society at large, to improve the human condition. Therefore, we need to research things that are relevant to the challenges facing humanity, like access to water and food, as well as affordable and clean energy.

In addition, we say this is a research university that does teaching, not the other way around. Every faculty member at KFUPM must engage in research, and we do this through more than 20 new interdisciplinary research centers created over the past 5 years. The crux of each of these research centers is interdisciplinary activities. We believe that knowledge is created at the intersection of disciplines, that it is the nexus of disciplines that gives rise to innovation. If you limit yourself to a certain discipline, you plateau very quickly and run out of ideas.

How does KFUPM's research tie into industry's needs and real-world global challenges?

Research cannot be done in a vacuum—you need to have a constant pollination of industry and academic thinking. The short answer is that we emphasize strong links with industry, and we strive to be a "tech-leaping" institution. That means that we should foresee challenges facing humanity and start developing and articulating solutions for them right now, before industry comes knocking. Hence, all of our 100 new programs are interdisciplinary, and all of them are future-facing. Our programs include topics such as AI and machine learning, cybersecurity and blockchain, visual computing, and quantum computing.

How do you nurture and encourage entrepreneurship among faculty and students?

Entrepreneurship is the natural catalyst that turns research into actual products. We are recognizing 2025 as the year of entrepreneurship at KFUPM, and we reorganized the entrepreneurship function at KFUPM from the ground up. For example, we tied the creation of new ventures to performance evaluations of our faculty. We have developed what we call our 0-0-0 policy to unlock full startup ownership rights for faculty and students, whether Saudi or international. Simply put, we require zero equity, zero royalty, and zero other tangential fees from our faculty and students, in order to support their success. It is now practically free for faculty and students to pursue their dream of establishing new startups with no strings attached—absolutely none.

Other highlights include a postdoc fellowship for entrepreneurs, where postdocs spend 50% of their time developing startups, and a new fellowship for undergraduate student founders. In addition, we are building a large prototyping center and scale-up facilities. We are launching the KFUPM Venture Craft Challenge, a worldwide global prize for startup ideas that have the potential to make a real impact on society.

AI is a key element of the educational program at KFUPM. Could you explain how AI is incorporated into the curriculum and why?

We have to foresee the future before industry develops that future, and AI is one concrete example of this. When we started the transformation 5 years ago, AI was not a hot technology, but we knew that AI is to this generation what mathematics was to our generation. You could not be an engineer, a scientist, or a financier if you didn't study mathematics. So, we came up with something that we called AI + X. You first learn AI at KFUPM, and then you learn X, X being your major. That is a requirement for all students, and it has added significant value to them in the marketplace because employers value this knowledge.

As the university continues to evolve, what are your aspirations for KFUPM and its place within the international academic and research landscape?

We want to transform KFUPM from a good regional university to a top international university whose impact is felt. I think becoming a top international university has, to some degree, been achieved, but I want the second part of that statement to be true. Our aspiration is that somebody will say, "Humanity is getting better access to clean water because of this technology that was invented at KFUPM." This is my personal aspiration and why I came to KFUPM—to make an impact on humanity. It will be a long journey, but we are in it till the end.

Sponsored by

KFUPM

PHOTO: PHOTOS PROVIDED BY KFUPM

Science AAAS

EDITOR-IN-CHIEF **Holden Thorp**, hthorp@aaas.org

EXECUTIVE EDITOR Valda Vinson

EDITORS, RESEARCH Sacha Vignieri, Jake S. Yeston **EDITOR, COMMENTARY** Lisa D. Chong

DEPUTY EXECUTIVE EDITOR Lauren Kmec

DEPUTY EDITORS Stella M. Hurtley (UK), Phillip D. Szuromi **SENIOR EDITORS** Caroline Ash (UK), Michael A. Funk, Angela Hessler, Di Jiang, Priscilla N. Kelly, Marc S. Lavine (Canada), Sarah Lempriere (UK), Mattia Maroso, Yevgeniya Nusinovich, Ian S. Osborne (UK), Joana Osório (UK), L. Bryan Ray, H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Yury V. Suleymanov, Valerie B. Thompson, Brad Wible **ASSOCIATE EDITORS** Jack Huang, Sumin Jin, Bianca Lopez, Sarah Ross (UK), Madeleine Seale (UK), Corinne Simonti, Ekeoma Uzogara **SENIOR LETTERS EDITOR** Jennifer Sills **NEWSLETTER EDITOR** Christie Wilcox **RESEARCH & DATA ANALYST** Jessica L. Slater **LEAD CONTENT PRODUCTION EDITORS** Chris Filiatreau, Harry Jach **SR. CONTENT PRODUCTION EDITOR** Amelia Beyna **CONTENT PRODUCTION EDITORS** Anne Abraham, Robert French, Julia Haber-Katris, Nida Masiulis, Abigail Shashikanth, Suzanne M. White **SENIOR PROGRAM ASSOCIATE** Maryrose Madrid **EDITORIAL MANAGER** Joi S. Granger **EDITORIAL ASSOCIATES** Aneera Dobbins, Lisa Johnson, Jerry Richardson, Anita Wynn **SENIOR EDITORIAL COORDINATORS** Alexander Kief, Ronmel Navas, Isabel Schnaidt, Alice Whaley (UK), Brian White **EDITORIAL COORDINATORS** Samuel Bates, Clair Goodhead (UK), Kat Kirkman, Daniel Young **ADMINISTRATIVE COORDINATOR** Karalee P. Rogers **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI OFFICE MANAGER** Carly Hayward (UK) **ASI SR. OFFICE ADMINISTRATORS** Simon Brignell (UK), Jessica Waldock (UK) **COMMUNICATIONS DIRECTOR** Meagan Phelan **DEPUTY DIRECTOR** Matthew Wright **SENIOR WRITERS** Walter Beckwith, Joseph Cariz, Abigail Eisenstadt **WRITER** Mahathi Ramaswamy **SENIOR COMMUNICATIONS ASSOCIATES** Zachary Graber, Sarah Woods **COMMUNICATIONS ASSOCIATES** Kiara Brooks, Haley Riley, Mackenzie Williams

NEWS EDITOR Tim Appenzeller

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** David Malakoff **DEPUTY NEWS EDITORS** Rachel Bernstein, Martin Enserink, David Grimm, Eric Hand, Michael Price, Kelly Servick, Matt Warren (Europe) **SENIOR CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis **ASSOCIATE EDITORS** Michael Greshko, Katie Langin **NEWS REPORTERS** Jeffrey Brainard, Adrian Cho, Phie Jacobs, Jocelyn Kaiser, Robert F. Service, Jennie Erin Smith, Erik Stokstad, Paul Voosen **CONSULTING EDITOR** Elizabeth Culotta **CONTRIBUTING CORRESPONDENTS** Vaishnavi Chandrashekhar, Dan Charles, Warren Cornwall, Andrew Curry (Berlin), Ann Gibbons, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Virginia Morell, Dennis Normile (Tokyo), Catherine Offord, Cathleen O'Grady, Elisabeth Pain (Careers), Charles Piller, Zack Savitsky, Richard Stone (Senior Asia Correspondent), Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **INTERNS** Celina Zhao **COPY EDITORS** Julia Cole (Senior Copy Editor), Hannah Knighton, Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

CREATIVE DIRECTOR Beth Rakouskas

DESIGN MANAGING EDITOR Chrystal Smith **GRAPHICS MANAGING EDITOR** Chris Bickel **PHOTOGRAPHY MANAGING EDITOR** Emily Petersen **MULTIMEDIA MANAGING PRODUCER** Kevin McLean **DIGITAL DIRECTOR** Kara Estelle-Powers **DESIGN EDITOR** Marcy Atarod **DESIGNER** Noelle Jessup **SENIOR SCIENTIFIC ILLUSTRATOR** Noelle Burgess **SCIENTIFIC ILLUSTRATORS** Austin Fisher, Kellie Holoski, Ashley Mastin **SENIOR GRAPHICS EDITOR** Monica Hersher **GRAPHICS EDITOR** Veronica Penney **SENIOR PHOTO EDITOR** Charles Borst **PHOTO EDITOR** Elizabeth Billman **SENIOR PODCAST PRODUCER** Sarah Crespi **SENIOR VIDEO PRODUCER** Meagan Cantwell **SOCIAL MEDIA STRATEGIST** Jessica Hubbard **SOCIAL MEDIA PRODUCER** Sabrina Jenkins **WEB DESIGNER** Jennie Pajerowski

CHIEF EXECUTIVE OFFICER AND EXECUTIVE PUBLISHER

Sudip Parikh

PUBLISHER, SCIENCE FAMILY OF JOURNALS Bill Moran

DIRECTOR, BUSINESS OPERATIONS & ANALYSIS Eric Knott **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **SENIOR MANAGER, BUSINESS ANALYSIS** Cory Lipman **BUSINESS ANALYSTS** Kurt Ennis, Maggie Clark, Isacco Fusi **BUSINESS OPERATIONS ADMINISTRATOR** Taylor Fisher **DIGITAL SPECIALIST** Marissa Zuckerman **SENIOR PRODUCTION MANAGER** Jason Hillman **SENIOR MANAGER, PUBLISHING AND CONTENT SYSTEMS** Marcus Spiegler **CONTENT OPERATIONS MANAGER** Rebecca Doshi **PUBLISHING PLATFORM MANAGER** Jessica Loayza **PUBLISHING SYSTEMS SPECIALIST, PROJECT COORDINATOR** Jacob Hedrick **SENIOR PRODUCTION SPECIALIST** Kristin Wowk **PRODUCTION SPECIALISTS** Kelsey Cartelli, Audrey Diggs **SPECIAL PROJECTS ASSOCIATE** Shantel Agnew

MARKETING DIRECTOR Sharice Collins **ASSOCIATE DIRECTOR, MARKETING** Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **ASSOCIATE DIRECTOR, MARKETING SYSTEMS & STRATEGY** Aimee Aponte **SENIOR MARKETING MANAGER** Shawana Arnold **MARKETING MANAGER** Ashley Evans **MARKETING ASSOCIATES** Hugues Beaulieu, Ashley Hylton, Lorena Chirinos Rodriguez, Jenna Voris **MARKETING ASSISTANT** Courtney Ford **SENIOR DESIGNER** Kim Huynh

DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING Erika Gebel Berg **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SENIOR TRAFFIC ASSOCIATE** Christine Hall

DIRECTOR, PRODUCT MANAGEMENT Kris Bishop **PRODUCT DEVELOPMENT MANAGER** Scott Chernoff **ASSOCIATE DIRECTOR, PUBLISHING INTELLIGENCE** Rasmus Andersen **SR. PRODUCT ASSOCIATE** Robert Koepke **PRODUCT ASSOCIATES** Caroline Breul, Anne Mason

ASSOCIATE DIRECTOR, INSTITUTIONAL LICENSING MARKETING Kess Knight **ASSOCIATE DIRECTOR, INSTITUTIONAL LICENSING SALES** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Nazim Mohammedi, Claudia Paulsen-Young **SENIOR MANAGER, INSTITUTIONAL LICENSING OPERATIONS** Judy Lillibridge **MANAGER, RENEWAL & RETENTION** Lana Guz **SYSTEMS & OPERATIONS ANALYST** Ben Teincuff **FULFILLMENT ANALYST** Aminta Reyes

ASSOCIATE DIRECTOR, INTERNATIONAL Roger Goncalves **ASSOCIATE DIRECTOR, US ADVERTISING** Stephanie O'Connor **US MID WEST, MID ATLANTIC AND SOUTH EAST SALES MANAGER** Chris Hoag **DIRECTOR, OUTREACH AND STRATEGIC PARTNERSHIPS, ASIA** Shoupeng Liu **SALES REP, ROW** Sarah Lelarge **SALES ADMIN ASSISTANT, ROW** Victoria Glasbey **DIRECTOR OF GLOBAL COLLABORATION AND ACADEMIC PUBLISHING RELATIONS, ASIA** Xiaoying Chu **ASSOCIATE DIRECTOR, INTERNATIONAL COLLABORATION** Grace Yao **SALES MANAGER** Danny Zhao **MARKETING MANAGER** Kilo Lan **ASCA CORPORATION, JAPAN** Rie Rambelli (Tokyo), Miyuki Tani (Osaka)

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **LICENSING ASSOCIATE** Virginia Warren **RIGHTS AND LICENSING COORDINATOR** Dana James **CONTRACT SUPPORT SPECIALIST** Michael Wheeler

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

INFORMATION FOR AUTHORS
science.org/authors/
science-information-authors

REPRINTS AND PERMISSIONS
science.org/help/
reprints-and-permissions

MULTIMEDIA CONTACTS
SciencePodcast@aaas.org
ScienceVideo@aaas.org

MEDIA CONTACTS
scipak@aaas.org

PRODUCT ADVERTISING & CUSTOM PUBLISHING
advertising.science.org
science_advertising@aaas.org

CLASSIFIED ADVERTISING
advertising.science.org/
science-careers
advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE
employers.sciencecareers.org
support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
science.org/subscriptions

MEMBER BENEFITS
aaas.org/membership/
benefits

INSTITUTIONAL SALES AND SITE LICENSES
science.org/librarian

AAAS BOARD OF DIRECTORS

CHAIR Joseph S. Francisco

IMMEDIATE PAST PRESIDENT
Willie E. May

PRESIDENT
Theresa A. Maldonado

PRESIDENT-ELECT
Marina Picciotto

TREASURER Carolyn N. Ainslie

COUNCIL CHAIR
Ichiro Nishimura

CHIEF EXECUTIVE OFFICER
Sudip Parikh

BOARD
Mark Dingfield
Morton Ann Gernsbacher
Kathleen Hall Jamieson
Jane Maienschein
Babak Parviz
Gabriela Popescu
Juan S. Ramírez Lugo
Susan M. Rosenberg
Vassiliki Betty Smocovitis
Roger Wakimoto

BOARD OF REVIEWING EDITORS (Statistics board members indicated with **(S)**)

Erin Adams, *U. of Chicago*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Fdn.*
Anastassia Alexandrova, *UCLA*
Mohammed AlQuraishi, *Columbia U.*
James Analytis, *UC Berkeley*
Paola Arlotta, *Harvard U.*
Madan Babu, *St. Jude*
Jennifer Balch, *U. of Colorado*
Nenad Ban, *ETH Zürich*
Carolina Barillas-Mury, *NIH, NIAID*
Christopher Barratt, *U. of Dundee*
François Barthelat, *U. of Colorado Boulder*
Franz Bauer, *Universidad de Tarapacá*
Andreas Baumler, *UC Davis*
Carlo Beenakker, *Leiden U.*
Sarah Bergbreiter, *Carnegie Mellon U.*
Kiros T. Berhane, *Columbia U.*
Aude Bernheim, *Inst. Pasteur*
Joseph J. Berry, *NREL*
Aditi Bhargav, *UCSF*
Dominique Bonnet, *Francis Crick Inst.*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Malcolm Brenner, *Baylor Coll. of Med.*
Ron Brookmeyer, *UCLA* **(S)**
Christian Büchel, *UKE Hamburg*
Johannes Buchner, *TUM*
Dennis Burton, *Scripps Res.*
Carter Tribley Butts, *UC Irvine*
György Buzsáki, *NYU School of Med.*
Annmarie Carlton, *UC Irvine*
Jane Carlton, *Johns Hopkins U.*
Pedro Carvalho, *U. of Oxford*
Simon Cauchemez, *Inst. Pasteur*
Ling-Ling Chen, *SIBCB, CAS*
Hilde Cheroutre, *La Jolla Inst.*
Wendy Cho, *UIUC*
Ib Chorkendorff, *Denmark TU*
Chunaram Choudhary, *Københavns U.*
Karlene Cimprich, *Stanford U.*
Laura Colgin, *UT Austin*
James J. Collins, *MIT*
Robert Cook-Deegan, *Arizona State U.*
Carolyn Coyne, *Duke U.*
Roberta Croce, *VU Amsterdam*
Ismaila Dabo, *Penn State U.*
Nicolas Dauphas, *U. of Chicago*
Claude Desplan, *NYU*
Sandra Díaz, *U. Nacional de Córdoba*
Samuel Díaz-Muñoz, *UC Davis*
Ulrike Diebold, *TU Wien*
Stefanie Dimmeler, *Goethe-U. Frankfurt*
Hong Ding, *Inst. of Physics, CAS*
Dennis Discher, *UPenn*
Jennifer A. Doudna, *UC Berkeley*
Ruth Drdla-Schutting, *Med. U. Vienna*
Raissa M. D'Souza, *UC Davis*
Bruce Dunn, *UCLA*
William Dunphy, *Caltech*
Scott Edwards, *Harvard U.*
Todd A. Ehlers, *U. of Glasgow*
Tobias Erb, *MPS, MPI Terrestrial Microbiology*
Beate Escher, *UFZ & U. of Tübingen*
Barry Everitt, *U. of Cambridge*
Vanessa Ezenwa, *U. of Georgia*
Toren Finkel, *U. of Pitt. Med. Ctr.*
Natascha Förster Schreiber, *MPI Extraterrestrial Phys.*
Serita Frey, *U. of New Hampshire*
Elaine Fuchs, *Rockefeller U.*
Caixia Gao, *Inst. of Genetics and Developmental Bio., CAS*
Daniel Geschwind, *UCLA*
Lindsey Gillson, *U. of Cape Town*
Alemu Gonsamo Gosa, *McMaster U.*
Simon Greenhill, *U. of Auckland*
Gillian Griffiths, *U. of Cambridge*
Nicolas Gruber, *ETH Zürich*
Hua Guo, *U. of New Mexico*
Taekjip Ha, *Johns Hopkins U.*
Daniel Haber, *Mass. General Hos..*
Hamida Hammad, *VIB IRC*
Brian Hare, *Duke U.*
Kelley Harris, *U. of Wash*
Carl-Philipp Heisenberg, *IST Austria*
Emiel Hensen, *TU Eindhoven*
Christoph Hess, *U. of Basel & U. of Cambridge*
Brian Hie, *Stanford U.*
Heather Hickman, *NIAID, NIH*
Janneke Hille Ris Lambers, *ETH Zürich*
Kai-Uwe Hinrichs, *U. of Bremen*
Pinshane Huang, *UIUC*
Christina Hulbe, *U. of Otago, New Zealand*
Gwyneth Ingram, *ENS Lyon*
Darrell Irvine, *Scripps Res.*
Erich Jarvis, *Rockefeller U.*
Peter Jonas, *IST Austria*
Sheena Josselyn, *U. of Toronto*
Matt Kaeberlein, *U. of Wash.*
Daniel Kammen, *UC Berkeley*
Kisuk Kang, *Seoul Nat. U.*
Vedika Khemani, *Stanford U.*
V. Narry Kim, *Seoul Nat. U.*
Nancy Knowlton, *Smithsonian*
Etienne Koechlin, *École Normale Supérieure*
LaShanda Korley, *U. of Delaware*
Paul Kubes, *U. of Calgary*
Deborah Kurrasch, *U of Calgary*
Laura Lackner, *Northwestern U.*
Mitchell A. Lazar, *UPenn*
Hedwig Lee, *Duke U.*
Fei Li, *Xi'an Jiaotong U.*
Jianyu Li, *McGill U.*
Ryan Lively, *Georgia Tech*
Luis Liz-Marzán, *CIC biomaGUNE*
Omar Lizardo, *UCLA*
Jonathan Losos, *WUSTL*
Ke Lu, *Inst. of Metal Res., CAS*
Christian Lüscher, *U. of Geneva*
Jean Lynch-Stieglitz, *Georgia Tech*
David Lyons, *U. of Edinburgh*
Fabienne Mackay, *QIMR Berghofer*
Vidya Madhavan, *UIUC*
Anne Magurran, *U. of St. Andrews*
Oscar Marín, *King's Coll. London*
Matthew Marinella, *Arizona State U.*
Charles Marshall, *UC Berkeley*
Christopher Marx, *U. of Idaho*
Geraldine Masson, *CNRS*
Jennifer McElwain, *Trinity College Dublin*
Scott McIntosh, *NCAR*
Rodrigo Medellín, *U. Nacional Autónoma de México*
Mayank Mehta, *UCLA*
C. Jessica Metcalf, *Princeton U.*
Tom Misteli, *NCI, NIH*
Jeffery Molkentin, *Cincinnati Children's Hospital Medical Center*
Alison Motsinger-Reif, *NIEHS, NIH* **(S)**
Rosa Moysés, *U. of São Paulo School of Medicine*
Carey Nadell, *Dartmouth College*
Thi Hoang Duong Nguyen, *MRC LMB*
Pilar Ossorio, *U. of Wisconsin*
Andrew Oswald, *U. of Warwick*
Isabella Pagano, *Istituto Nazionale di Astrofisica*
Giovanni Parmigiani, *Dana-Farber* **(S)**
Zak Page, *UT Austin*
Sergiu Pasca, *Stanford U.*
Pierre Paoletti, *Ecole Normale Supérieure, Paris*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Philip Phillips, *UIUC*
Matthieu Piel, *Inst. Curie*
Kathrin Plath, *UCLA*
Martin Plenio, *Ulm U.*
Katherine Pollard, *UCSF*
Elvira Poloczanska, *Alfred-Wegener-Inst.*
Julia Pongratz, *Ludwig Maximilians U.*
Philippe Poulin, *CNRS*
Suzie Pun, *U. of Wash*
Lei Stanley Qi, *Stanford U.*
Simona Radutoiu, *Aarhus U.*
Maanasa Raghavan, *U. of Chicago*
Trevor Robbins, *U. of Cambridge*
Adrienne Roeder, *Cornell U.*
Joeri Rogelj, *Imperial Coll. London*
John Rubenstein, *SickKids*
Sylvie Roke, *École Polytechnique Fédérale de Lausanne*
Yvette Running Horse Collin, *Toulouse U.*
Mike Ryan, *UT Austin*
Alberto Salleo, *Stanford U.*
Miquel Salmeron, *Lawrence Berkeley Nat. Lab*
Nitin Samarth, *Penn State U.*
Erica Ollmann Saphire, *La Jolla Inst.*
Joachim Saur, *U. zu Köln*
Alexander Schier, *Harvard U.*
Wolfram Schlenker, *Columbia U.*
Susannah Scott, *UC Santa Barbara*
Anuj Shah, *U. of Chicago*
Vladimir Shalaev, *Purdue U.*
Jie Shan, *Cornell U.*
Jay Shendure, *U. of Wash.*
Steve Sherwood, *U. of New South Wales*
Ken Shirasu, *RIKEN CSRS*
Robert Siliciano, *JHU School of Med.*
Emma Slack, *ETH Zürich & U. of Oxford*
Richard Smith, *UNC* **(S)**
Ivan Soltesz, *Stanford U.*
John Speakman, *U. of Aberdeen*
Allan C. Spradling, *Carnegie Institution for Sci.*
V. S. Subrahmanian, *Northwestern U.*
Sandip Sukhtankar, *U. of Virginia*
Naomi Tague, *UC Santa Barbara*
A. Alec Talin, *Sandia Natl. Labs*
Patrick Tan, *Duke-NUS Med. School*
Sarah Teichmann, *Wellcome Sanger Inst.*
Dörthe Tetzlaff, *Leibniz Institute of Freshwater Ecology and Inland Fisheries*
Amanda Thomas, *UC Davis*
Rocio Titiunik, *Princeton U.*
Shubha Tole, *Tata Inst. of Fundamental Res.*
Maria-Elena Torres Padilla, *Helmholtz Zentrum München*
Kimani Toussaint, *Brown U.*
Barbara Treutlein, *ETH Zürich*
Li-Huei Tsai, *MIT*
Jason Tylianakis, *U. of Canterbury*
Matthew Vander Heiden, *MIT*
Jo Van Ginderachter, *VIB, U. of Ghent*
Ivo Vankelecom, *KU Leuven*
Henrique Veiga-Fernandes, *Champalimaud Fdn.*
Reinhilde Veugelers, *KU Leuven*
Elizabeth Villa, *UC San Diego*
Bert Vogelstein, *Johns Hopkins U.*
David Wallach, *Weizmann Inst.*
Jane-Ling Wang, *UC Davis* **(S)**
Jessica Ware, *Amer. Mus. of Natural Hist.*
David Waxman, *Fudan U.*
Alex Webb, *U. of Cambridge*
Chris Wikle, *U. of Missouri* **(S)**
Ian A. Wilson, *Scripps Res.* **(S)**
Marjorie Weber, *U. of Michigan*
Sylvia Wirth, *ISC Marc Jeanneroд*
Hao Wu, *Harvard U.*
Amir Yacoby, *Harvard U.*
Benjamin Youngblood, *St. Jude*
Yu Xie, *Princeton U.*
Kenneth Zaret, *UPenn School of Med.*
Lidong Zhao, *Beihang U.*
Bing Zhu, *Inst. of Biophysics, CAS*
Xiaowei Zhuang, *Harvard U.*
Maria Zuber, *MIT*

Saving science by the sea

Susan Fitzpatrick and **Jane Maienschein**

As funding for science tightens across the United States, attention has turned to pressures faced by universities and biomedical research institutions. An often overlooked part of the nation's science ecosystem, however, are the small, independent marine laboratories that are also bearing the brunt of funding cuts and other constraints imposed by the Trump administration. These institutions have long nurtured generations of scientists—from high school students to Nobel laureates—and precipitated discoveries that have advanced modern science. Yet they now face an uncertain future, and losing them would eliminate not only a powerful convergence of research and education but also a wellspring of scientific advancement.

Marine laboratories occupy a rare niche. They are mostly independent but often loosely affiliated with universities by sharing faculty, space, or other resources. They host students and scientists from weeks to months, mixing disciplines and experience levels that break down traditional academic silos across the spectrum of life sciences, from molecular to microbial to ecological. The knowledge that many of the participants will be there for only a short period creates a sense of urgency for thinking beyond the typical and asking questions that challenge assumptions and lead to exploring problems in new ways. People come to work intensively in this atmosphere and in nonheirarchical teams, sharing ideas over meals and debating hard questions long into the night. They work together, live together, and share the excitement of discovery together.

These laboratories are... an important scientific culture.

The Marine Biological Laboratory in Woods Hole, Massachusetts, founded in 1888, established the precedent. Forged by a Boston consortium committed to educating both men and women, it became the model for bringing together diverse perspectives. It also opened the research community to new environments and organisms as model systems for studying a wide scope of problems, from the details of embryonic development to the scope of biodiversity in different habitats. As physician and writer Lewis Thomas put it, this "national laboratory" for biology serves as a paradigm for collaborative integrated investigation and learning. It was here in 1954 that graduate students Matthew Meselson and Franklin Stahl met and began a collaboration that led to their elegant experiment showing how DNA replicates.

There are more than 100 marine research institutions in the United States that hold memberships in the National Association of Marine Laboratories (NAML) (plus others that are not members). This includes university-affiliated and independent laboratories, many as nonprofit entities. Some are well known for their emphasis on oceans (The Scripps Institution of Oceanography and Bermuda Institute of Ocean Sciences among them) or on integrating research with education (Cold Spring Harbor Laboratory and MDI Biological Laboratory, for example). Each has its own focus, but all promote a tight-knit community driven by curiosity, an immersive experience in the natural environment of their science, and a spirit of openness to trying new things and being willing to fail.

Collectively, these marine laboratories have trained hundreds of thousands of students and researchers, many of whom have gone on to careers in academia, government, industry, and other sectors. Yet, the laboratories' independence, which has been a strength, may now be a vulnerability with funding mechanisms that favor large research institutions. These modest laboratories cannot adjust as readily to new constraints on funding and on more risk-taking research. For example, some are reporting the need to fire staff, curtail student internships, cut back on course offerings, and take other actions that affect both research and training of the next generation of scientists. NAML was established in 1989 in response to funding cuts in 1984 for marine research administered by the National Oceanic and Atmospheric Administration. Marine laboratory directors and scientists joined together to forge a forum to address common challenges across the marine science enterprise. Recent NAML meetings, including earlier this month, continue to address the future of federal science partners and relevant funding and policy issues to strengthen the marine science community against vulnerabilities and maximize opportunities for the field.

Marine laboratories encourage exploratory and creative thinking. They are a reminder that progress is often the result of serendipity—through conversations in dining halls, working side by side in a shared laboratory space, or a chance meeting during a coffee break. These laboratories are not just generators of data but incubators of an important scientific culture. Against the current backdrop of anxiety about the future of science, the scientific community and public need these places now more than ever. □

Susan Fitzpatrick is president emerita of the James S. McDonnell Foundation, St. Louis, MO, USA, and a trustee of the Marine Biological Laboratory, Woods Hole, MA, USA. susan@jsmf.org **Jane Maienschein** is university professor in the School of Life Sciences at Arizona State University, Tempe, AZ, USA, and a fellow of the Marine Biological Laboratory, Woods Hole, MA, USA. maienschein@asu.edu

10.1126/science.aed1909



A decade in the making, the China-Argentina Radio Telescope in the foothills of the Andes is nearly complete.

PHOTO: RICARDO CÉSAR PODESTÁ

ASTRONOMY

Argentine telescope caught in U.S.-China crossfire

Superpower tensions have abruptly derailed a nearly finished radio telescope **MARÍA DE LOS ÁNGELES ORFILA**

Later this year, astronomers in Argentina expected to begin testing the largest radio dish in Latin America: a 40-meter-diameter instrument designed to detect electromagnetic whispers from the distant cosmos. The Argentine and Chinese governments launched the project a decade ago as a symbol of scientific cooperation between the two nations. But now the China-Argentina Radio Telescope (CART), which is nearly finished, is stuck in limbo—a victim of growing tensions between the United States and China.

"We're 90% complete. Very close. That's why what's happening is so difficult, because we don't want it to become scrap metal," says María Verónica Benavente, a senior official at the National University of San Juan (UNSJ), which is leading the project.

Researchers trace CART's difficulties to efforts by Argentina's libertarian president, Javier Milei, to forge close ties with U.S. President Donald Trump, whose administration has made curbing China's global influence a priority. The first signs of trouble came in June, when Argentina's science agency missed a deadline for renewing a scientific cooperation agreement with China. Then, in September, customs officials in Buenos Aires held up a shipment of key dish parts made in China, which has contributed some $15 million in equipment and engineering to the project. Last month brought another setback when U.S. Secretary of the Treasury Scott Bessent said the Trump administration would provide up to $20 billion to shore up Argentina's beleaguered economy—but only if Milei's government suspended strategic ties with China, including efforts to build "observational facilities."

That policy was aimed "directly" at CART, believes Benavente, who is the former science secretary for San Juan, the mountainous western province where the observatory is located.

"There is direct interference from the White House here."

The Chinese embassy in Argentina condemned Bessent's comments, saying they represented a "Cold War mentality." In a second statement, it warned that "Latin America and the Caribbean are no one's backyard."

The latest setback came on 17 October, when Milei's government issued a requirement that the Ministry of Defense approve "any installation of radars, observatories, or aerospace systems." That rule is especially vexing because "CART is 100% scientific," says Ricardo Podestá, director of the Félix Aguilar Astronomical Observatory, where the radio telescope is being built. "If this were a military project, everyone would notice."

Others note that the requirement seems redundant because CART already received Defense Ministry approval in 2016 under a previous president who is politically aligned with Milei. "I don't know what else this government wants," says Jorge Castro, a science dean at UNSJ. "I hope they're simply unaware of this objective fact."

The Milei administration's policies are "simply an over-the-top move to show themselves obsequiously aligned with the Trump administration, to the detriment of the entire astronomical community in Argentina," Roberto Salvarezza, head of the Buenos Aires Scientific Research Commission, told the media outlet *Página/12*.

The delays have left astronomers frustrated. "If the equipment is released, we could have the antenna ready this year or early next year," Podestá says, opening the way to a wide range of research, including studies of celestial objects such as pulsars and quasars, as well as collaborations with the more than 100 other radio telescopes around the world. "San Juan has some of the clearest skies in the world," Podestá notes. "You can see the galactic nucleus high in the sky."

To ensure that such observing opportunities don't go to waste, the government should issue an "immediate guarantee of [CART's] continuity," UNSJ's science faculty said in a statement. "The development of science has no borders," it added, "and a cosmopolitan vision must prevail." □

María de los Ángeles Orfila is a journalist in Uruguay.

SOCIAL MEDIA RESEARCH

Europe faults Big Tech platforms for lack of research access

Scientists welcome preliminary finding that Meta and TikTok are in breach of the Digital Services Act **KAI KUPFERSCHMIDT**

When Philipp Lorenz-Spreen set out in 2024 to study how politicians across Europe communicate online and how much divisive language they use, he knew he had the law on his side. The European Union's Digital Services Act (DSA), which had come into force in February of that year, guaranteed researchers like Lorenz-Spreen, a computational social scientist at the Dresden University of Technology, access to data from social media platforms X, TikTok, Facebook, and Instagram. All he had to do was ask.

But the reality was frustrating. X kept asking follow-up questions and eventually denied his request without giving a reason. TikTok provided access, but Lorenz-Spreen decided not to rely on the company because other researchers had found big gaps in the data it shared. Meta, the parent company of Facebook and Instagram, offered an online dashboard that came with serious limitations; for example, it captured data only on accounts with more than 25,000 followers, more than some European politicians have. Lorenz-Spreen's conclusion: The DSA wasn't working the way it should.

> **"The public is entitled to know what platforms and their algorithms are doing to the information landscape."**
>
> **Stephen Lewandowsky**
> University of Bristol

Now, the European Commission has made clear it agrees. On 24 October, it announced that after an investigation it had "preliminarily found both TikTok and Meta in breach of their obligation to grant researchers adequate access to public data" under the DSA. (The same announcement also faulted Meta for its content moderation procedures.)

The Commission did not reveal details about its investigation nor say what prompted it. But the move, which follows a similar preliminary finding against X in July 2024, is a welcome sign the Commission is serious about enforcing the DSA, says Stephen Lewandowsky, a psychologist at the University of Bristol. "Access to data by researchers is a crucial element of ensuring transparency and democratic accountability," Lewandowsky says. "The public is entitled to know what platforms and their algorithms are doing to the information landscape."

The Commission can impose a billion-dollar fine on each company if it ultimately decides they are not complying with the DSA, but first the companies will have a chance to defend themselves. Rebekah Tromble, a political scientist at Northeastern University, says they might do so vigorously. "In recent months, many of these companies have been more aggressively pushing back" against the DSA's provisions, she says. In an email, a Meta spokesperson told *Science* the company was "confident" it had already taken the necessary steps to comply with the DSA. A spokesperson for TikTok wrote that the company is reviewing the Commission's findings and argued the DSA's data sharing requirements were in tension with EU privacy regulations.

Data from Big Tech is vital for researchers who want to study, say, how election lies spread on Facebook, or whether the algorithm determining what you see on TikTok leads to polarization. The DSA, enacted in 2022, guaranteed access for researchers—anywhere in the world—who are studying "systemic risks" for the EU, such as negative effects on elections or public health. At the time, only a few companies were relatively generous with such information. X, then still named Twitter, was seen as very transparent and as a result was "disproportionately studied," says Daphne Keller, an internet law expert at Stanford University. Other companies cooperated with small groups of scientists on research projects.

But "it is clear to me that those days are over," Tromble says. The rise of artificial intelligence has made things worse by

making companies more protective of their data, to keep them from being used to train large language models, says Przemyslaw Grabowicz, a researcher at University College Dublin.

LK Seiling of the Weizenbaum Institute, founded in Germany in 2017 to study the "networked society," and his colleagues have surveyed researchers about their experiences using the DSA. They found the platforms use several techniques to delay or deny requests for data. X, for example, has a very simple online application form that never seems to lead to quick acceptance. "If I'm lucky I'm rejected after 43 days, but if I'm not, then I end up in what I call the circle of hell," a loop of more and more questions that may or may not result in data, Seiling says. Even if they grant access, companies often make insufficient data available.

So far, researchers have only been able to request "publicly available" information under the DSA—for instance, what specific users post or repost, or whom they follow. For a host of logistical and legal reasons, researchers often prefer the companies to provide such data, but they can also collect data themselves with the help of computer programs that "scrape" web pages and collect vast amounts of data. Some companies' terms of service ban scraping, but the DSA allows it, Keller says. (Many researchers are already using this technique but prefer not to talk about it.)

On 29 October, however, the DSA entered a new, likely even more contentious phase, in which researchers can also request data that aren't public, for example what content is shown to individual people. The platforms are expected to protect such information even more zealously, but for researchers, it's an opportunity to better understand how they really work, Lorenz-Spreen says.

Take, for example, the well-documented finding that outrage-inducing content gets more likes. "The question always remains whether that is because that's people's psychology, and they just like this content more, or whether it is somehow being pushed by the platforms," Lorenz-Spreen says. "That kind of data is what we really need if we want to make serious progress in our field."

The investigation that led to a preliminary finding against X last year is still continuing, a Commission spokesperson says. "This is the first case we opened under the DSA, and we are building the case to make sure it is solid."

Researchers and governments elsewhere will be watching Europe's battles closely. U.S. President Donald Trump's administration has attacked misinformation research as a gateway to censorship, and a Republican-led committee in the U.S. House of Representatives called the DSA "the EU's comprehensive digital censorship law," silencing conservative voices and infringing on "American free speech." U.S. Secretary of State Marco Rubio has reportedly ordered U.S. diplomats to lobby against the DSA.

That makes last week's finding even more important, Grabowicz says: "It signals that the Commission is willing to resist the pressure from right-wing U.S. policymakers." □

BY THE NUMBERS

40%

Proportion of more than 240 papers on animal models of hemorrhagic stroke found to have problematic or duplicated images in a new analysis. (*PLOS Biology*)

BIOMEDICINE

New antivenom is a 'potential game changer' for snakebites

Synthetic cocktail of llama and alpaca "nanobodies" protected mice **CHRISTIE WILCOX**

Venomous snakebites kill upward of 100,000 people every year and maim many times more, a toll that has led the World Health Organization to call snakebite a neglected tropical disease. Current antivenoms have done all they can do to stem it. But a cocktail of lab-synthesized "nanobodies" inspired by alpacas and llamas could change the picture. They should be safer to administer, cheaper to produce, and more effective than the best available antivenoms, according to a report this week in *Nature*.

This new kind of antivenom has proved protective only in mice so far. But "the science is solid," says Robert Rono, a doctor and epidemiologist with the Kenyan Ministry of Health in Baringo County, where snakebite is a pernicious public health issue. "It's a potential game changer if all goes well."

Current antivenoms can save lives and limbs if administered quickly and properly, but people in need rarely get them. "In our part of the world, a bite from a black mamba or cobra is almost akin to a death sentence," Rono says. "It's a very unfair disease that really hits the poorest of the poor," adds Andreas Laustsen-Kiel, a toxinologist and bioengineer at the Technical University of Denmark whose lab produced the new antivenom.

Infrastructure is one problem: Most bites occur in rural areas, where clinics are small. Even if a person reaches a larger, regional facility, there's no guarantee the doctors there will have enough antivenom available or the confidence in the drugs to use them.

Current products are also a barrier. They are made by injecting animals, usually horses or sheep, with snake venoms and harvesting the antibodies. The care and upkeep of these animals is expensive, which drives up the price. And because the mixtures are a mishmash of antibodies, some effective and others not, a lot of antivenom is often required to treat a bite. A regional center might have just one or two vials—and a single bitten person could need up to 10, Rono says.

Furthermore, the foreign antibodies sometimes set off a strong immune reaction, known



An innovative antivenom protected mice from lethal doses of venom from 17 of 18 African snakes—all except the eastern green mamba (*Dendroaspis angusticeps*).

PHOTO: WOLFGANG WÜSTER

as anaphylactic shock, that is also life-threatening. Many medical professionals in Africa are hesitant to administer the drugs, having had bad experiences with them.

To make their alternative, Laustsen-Kiel and colleagues injected mixtures of venoms from 18 African snakes into llamas and alpacas. Their antibodies are smaller than those of horses and humans and have only one protein chain, instead of two, on each arm of the Y-shaped molecule. Each arm, dubbed a nanobody, can be separated from its base while remaining able to bind a target molecule, such as a toxin. The isolated arms lack the animal-specific sections that set off allergic reactions.

The team took the DNA sequences for the llama and alpaca nanobodies and added them to *Escherichia coli* bacterial cells to mass produce them in vitro. Then they screened for the ones that most effectively inactivated the toxins of the 18 snakes, including 13 species of cobras, four species of mamba, and one rinkhal.

They ultimately identified eight nanobodies that, injected together with lethal doses of the venoms of the 18 snakes, protected mice from all but one, the eastern green mamba (*Dendroaspis angusticeps*). "You can take surprisingly few nanobodies … and actually get extremely broad coverage," Laustsen-Kiel says. When the mice were injected first with the venom and received the cocktail 5 minutes later, it protected them from eight of 11 snake species tested. Laustsen-Kiel says the group now plans to investigate adding more effective nanobodies for the toxins of the three remaining snakes.

For all species but the black mamba, the nanobody mix outperformed the currently available antivenom, Inoserp PAN-AFRICA. In addition to saving more mice, the cocktail prevented more tissue damage than PAN-AFRICA, notes Kartik Sunagar, an evolutionary biologist and toxinologist with the Indian Institute of Science, Bangalore. Laustsen-Kiel suggests the smaller nanobodies may make their way deeper into tissues to reach the wound site more effectively.

Rono believes that if medical professionals don't fear an allergic reaction, they will be more comfortable administering the nanobodies outside of large hospital settings, even in the fields or forests where people are bitten. "If the technology does what it says in the paper on the ground … this should be a very popular product," he says.

It might also be cheaper. Laustsen-Kiel predicts that synthesizing the nanobodies in the lab will be cheaper than procuring antibodies from live animals. And if it does end up being more effective per vial than current antivenoms, that should lower cost, too, Rono says.

Between tweaking the mixture and additional testing, Laustsen-Kiel estimates the team is about 3 years away from being ready for human testing. That will be the real hurdle, Sunagar says. "These formulations will need to successfully navigate human clinical trials." □

IN OTHER NEWS

ALZHEIMER'S FRAUD CASE DROPPED

A much-anticipated scientific fraud case against a neuroscientist for his preclinical work on a potential drug for Alzheimer's disease evaporated last week after the jury had been chosen, when the U.S. Department of Justice (DOJ) abruptly dropped the charges. Former City University of New York (CUNY) researcher Hoau-Yan Wang had been indicted in 2023 for "defrauding the National Institutes of Health of approximately $16 million" in grants related to early studies on a candidate Alzheimer's drug called simufilam. Its developer, biotech Cassava Sciences, abandoned the drug after it failed in a large, phase 3 trial. Federal prosecutors did not reply to questions about why they dropped the charges. Wang's defense attorneys had filed a request to dismiss the case, in part because they said DOJ failed to disclose a 30 May letter from CUNY concluding that Wang had not engaged in proven scientific misconduct, although it also noted that an investigative panel had found signs of improper image manipulation in scores of his studies. Wang's lawyers confirmed he has retired from CUNY.
—John Travis



POLAR SCIENCE

How NSF floats Antarctic research with no icebreaker

Ending *Palmer* lease is one of many belt-tightening moves amid budget uncertainty **JEFFREY MERVIS**

Palmer Station is the base of operations for many U.S. scientists studying the changing Antarctic ecosystem.

In 2022, the National Science Foundation (NSF) gave University of Alabama geoscientist Tom Tobin a grant to drill rock cores at a remote site off the Antarctic peninsula. NSF told Tobin, who works on mass extinctions, that he'd have to wait 4 years for a vessel to take him there.

This summer, however, NSF canceled the lease on his long-awaited ride, the *RV Nathaniel B. Palmer*, the only U.S. research icebreaker dedicated to working in the treacherous waters of the Southern Ocean. NSF has since rebooked him on another ship, but it's one that normally operates in the Arctic and lacks the *Palmer*'s capacity to crunch through thick sea ice.

Tobin's odyssey reflects the precarious state of the U.S. Antarctic Program (USAP), which NSF manages. Losing the *Palmer*, which NSF said it could no longer afford, comes only a year after NSF ended its lease for the *ARSV Laurence Gould*, which did double duty as a supply ship to NSF's Palmer Station as well as hosting research cruises. The belt-tightening in its $559 million Office of Polar Programs, which supports science and logistics in both the Arctic and Antarctic, has already led to downsizing and cancellations of long-running programs. Many more would be imperiled under President Donald Trump's plan to slash NSF's overall $9 billion budget to $4 billion in 2026.

Antarctic scientists say abandoning the *Palmer*, a mainstay of Antarctic research since 1992, is a big step in the wrong direction. "Access is the biggest barrier to doing Antarctic science, and losing the *Palmer* will mean less access," says Oscar Schofield, a biological oceanographer at Rutgers University who leads a long-running NSF-funded ecological monitoring project on the Antarctic Peninsula that was also scheduled to sail on the *Palmer* this winter. Instead, the *Palmer* completed its last cruise for NSF this month and reverted to its owner, Offshore Service Vessels.

NSF, which is currently closed because of the government shutdown, said in a 9 September posting it "plans to continue supporting the entire portfolio of marine cruise projects planned for the 2025-2026 field season … using the U.S. Academic Research Fleet, supported by University-National Oceanographic Laboratory System (UNOLS) operators."

The agency's decision to turn to UNOLS's 17-vessel fleet came as a surprise to Doug Russell, executive secretary for UNOLS. "I had no inkling of this," he says. Russell says NSF officials told him they needed to use UNOLS vessels because "we have important science programs to be covered down in Antarctica."

So far, NSF has kept its word by finding alternative rides for scien-

PHOTO: NATASJA VAN GESTEL/TEXAS TECH UNIVERSITY

tists who expected to be aboard the *Palmer*. And polar scientists seem willing to cut NSF some slack. "I think [canceling the *Palmer* lease] was shortsighted," says Julia Wellner, a marine geologist at the University of Houston who helped draft a 28 July petition signed by more than 200 scientists asking NSF to rethink its decision. "But NSF is caught between a rock and a hard place, and they're doing their best to keep the program going."

Even so, those new arrangements are only a short-term solution. Unlike the *Palmer*, which was based in Punta Arenas, Chile, the replacement ships have home ports thousands of kilometers away, forcing long voyages at daily operating rates of $75,000 or higher. And the ships are not available for other cruises during those weeks at sea.

In addition, the two ships that have so far taken on work scheduled for the *Palmer*—the *RV Sikuliaq* and the *RV Roger Revelle*—are in many ways less well suited for Antarctic cruises. The *Sikuliaq* has fewer berths for scientists, a shorter cruise range, and a less robust ice-breaking capacity than the larger *Palmer*. And although the *Revelle* is similar in size to the *Palmer*, it has never operated as far south as Antarctica and will have to maneuver around any ice hazards. Researchers aboard both ships might be forced to revise or abandon some expedition objectives if they encounter heavy ice.

Given such constraints, is the ship shuffling "really cheaper than maintaining an Antarctic research presence?" wonders Margaret Mars Brisbin, a biological oceanographer at the University of South Florida who, like Tobin, had her cruise switched to the *Sikuliaq*. "Sharing ships is going to really diminish the capacity of the U.S. polar program in general and make us less competitive with the rest of the world," she adds.

The upheaval has taken a toll, she and others say. NSF initially told Mars Brisbin that her upcoming cruise to measure how changing ice melt patterns affect phytoplankton growth and carbon emissions from the ocean was canceled and that she might even have to resubmit her proposal. When she got the good news that she and her colleagues would be using the *Sikuliaq*, she also learned they would depart much earlier than planned. They had to race to meet a 2-week deadline for a long list of logistical, medical, and paperwork requirements that usually take months to complete.

"Our cruise was going to be a month later than [Schofield's] and we were going to use some of their equipment," Mars Brisbin says. "But now they are happening at the same time, so we've had to hunt for replacements."

Schofield also had to scramble. "At first NSF put us on the *Sikuliaq*," Schofield says. "But it would have meant fewer people and a shorter cruise than the 8 weeks from dock to dock. It was also later in the season, which would have skewed the time series we've been collecting. So NSF decided to go with the *Revelle*, which is a much better fit for us."

Although Wellner applauds NSF for finding other ships for her colleagues, she worries their science could suffer because of the rushed preparations. "We think about every little thing on a ship, like what type of winch will be deploying our gear over the side," she says.

In addition to ships, many Antarctic scientists rely on remote sensing and satellite data from an NSF-funded Polar Geospatial Center at the University of Minnesota to pinpoint and collect data on their research locations. But this year the center, which provides such information free to any USAP-funded researcher, felt the ripple effects of NSF's looming budget crisis.

In May, after its application for a 5-year renewal had been stalled for more than a year, the center suspended accepting new requests for data. Shortly after, it laid off three of its 11 staffers. In August, the center received funding for another 5 years, but at roughly two-thirds of its previous operating level.

Ironically, some Arctic scientists wonder whether NSF's efforts to keep Antarctic science afloat are coming at their expense. Audrey Taylor, executive director of the Arctic Research Consortium of the United States (ARCUS), says assigning double duty to the *Sikuliaq* suggests Antarctica will get more than its share of whatever NSF allocates for polar science. ARCUS, whose $9 million NSF grant to build scientific networks and strengthen ties with Indigenous communities expires in early 2026, was preparing to compete for a new NSF-funded Arctic communications hub. But earlier this year, NSF decided not to fund the hub. Without any prospect of NSF support, ARCUS folded its tent on 30 September.

For his part, Tobin is content to have a way to get to his research site on Seymour Island, which offers an easily accessible fossil record of climate shifts following the impact of the meteor that wiped out the dinosaurs 65 million years ago. Although he's made four previous trips to Antarctica aboard the *Palmer* and the *Gould*, Tobin also carries out summer digs in Wyoming and Montana. "I could never spend all of my time doing polar research," he says, "because you never know when you'll be able to go down there." □



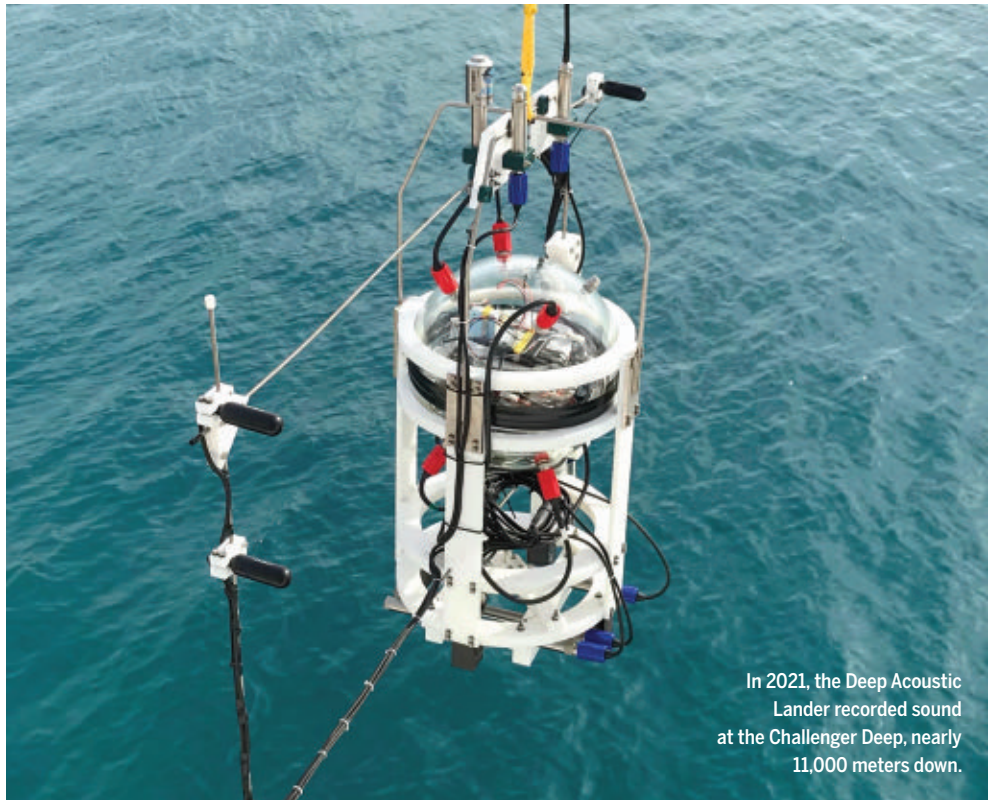
The National Science Foundation has ended the lease of its sole icebreaker, the *RV Nathaniel B. Palmer*.

PHOTO: NATASJA VAN GESTEL/TEXAS TECH UNIVERSITY

IN OTHER NEWS

PARATYPHOID VACCINE SHOWS ITS METTLE

A candidate vaccine against paratyphoid fever, a bacterial illness estimated to kill more than 20,000 people each year worldwide, has shown promise in a human "challenge" study, in which healthy volunteers are deliberately exposed to an infectious agent under tightly controlled circumstances. *Salmonella Paratyphi A*, which causes most cases of the illness, spreads via contaminated water or food, primarily in places that lack clean drinking water and sanitation. It can cause high fever, headache, abdominal pain, and diarrhea. In a trial at the University of Oxford, 70 volunteers were given either two doses of the vaccine—consisting of live but attenuated versions of the bacterium—or a placebo. Four weeks later, subjects in both groups swallowed a small amount of liquid containing bacteria from a wild *S. Paratyphi* strain isolated from a patient in Nepal in 2006. Compared with the placebo group, the vaccine reduced infection risk by 73%, researchers report this week in *The New England Journal of Medicine*. (Subjects in both groups who fell sick were cured with an antibiotic.) If licensed, the vaccine would likely be used in combination with existing oral vaccines for *Salmonella typhi*, which causes the more common—and more lethal—disease typhoid fever.
—Martin Enserink



In 2021, the Deep Acoustic Lander recorded sound at the Challenger Deep, nearly 11,000 meters down.

PHOTO: DAVID BARCLAY

OCEANOGRAPHY

Dangerous ocean acidification tracked using ambient noise

Acoustic technique could make it easier to monitor threat to marine life stemming from rising carbon emissions

PAUL VOOSEN

The ocean is a noisy place. Ship propellers and whale songs reverberate at the lowest pitches, while at higher tones dolphins click and shrimp snap their claws. Between these frequencies are the sounds of the churning sea itself, generated as waves, wind, and rain roil its surface. Researchers have now used this ambient noise to probe the rising acidity of the ocean. The acoustic technique, published this month in the *Journal of Geophysical Research: Oceans*, could make it easier to measure this key parameter of ocean health across vast distances rather than relying on point measurements.

"This is a really cool idea," says Andone Lavery, an acoustical oceanographer at Oregon State University. "And they've shown that it is possible."

The carbon emissions that are warming the globe also acidify seawater. The ocean naturally absorbs about one-third of annual carbon dioxide emissions; as this gas dissolves and reacts, it creates bicarbonate and hydrogen ions. The hydrogen lowers the pH of seawater, increasing ocean acidity, which can harm sea life and slow future carbon uptake. Ship-based measurements in shallow parts of the ocean have found that, since 1985, pH has already dropped from 8.11 to 8.04.

The waters of the ocean are layered, however, and measurements at one depth may not apply to others. Adding pH sensors to the thousands of robotic Argo floats that patrol the seas, diving as deep as 2000 meters, is one way to get a broader picture of acidity. But David Barclay, an acoustical oceanographer at Dalhousie University, and his co-authors found a way to measure average pH across even greater depth ranges, by taking advantage of the intrinsic physics of sound.

The work originated in 2009, when Barclay led a cruise to the Philippine Sea, funded by the U.S. Navy, to test a new acoustic listening device, called Deep Sound. Barclay's team dropped the instrument—two hydrophones and

their associated electronics encased in reinforced glass spheres—down 5000 meters. And then they listened.

One immediate surprise was that the abyss wasn't as quiet as they expected. They found that when surface winds whipped up to more than 10 knots or so, the sounds of breaking waves reached all the way to Deep Sound, at frequencies of about 1 to 10 kilohertz (kHz). Barclay knew that two compounds in the ocean, borate and magnesium sulfate, can dampen sounds at 1 kHz and 10 kHz, respectively, because of the particular way the molecules absorb a bit of the sound wave's energy.

Crucially, rising acidity decreases the abundance of borate, and its attenuating effect, while leaving magnesium sulfate unscathed. By comparing the two frequencies, the researchers thought they could get at the integrated pH of the entire water column, much as the echoes of earthquakes rippling through the ocean can be used to measure overall ocean heat.

It took 15 years to be sure: five more ship deployments and several upgrades of the instrument that ultimately allowed it to survive at depths of more than 10,000 meters in the Challenger Deep, in the Mariana Trench near the island of Guam. It also took a lot of painstaking mathematical analysis, says Ernst Uzhansky, an acoustic physicist now at the Naval Postgraduate School. "I can't tell you how happy I was," says Uzhansky, who was skeptical the team would find the signal. "Now it really opens up an opportunity to do large-scale pH monitoring."

Not everyone is convinced. Liqing Jiang, a chemical oceanographer at the University of Maryland, notes that the team did not compare its measurement with known observations. "This is a serious flaw," he says. The technique is not as accurate as measurements from the new biogeochemical (BGC) Argo probes, some 500 of which will finish being deployed by the United States by late next year. "The data shown in the paper is not the kind of quality we need yet," says Ken Johnson, a chemical oceanographer at the Monterey Bay Aquarium Research Institute and a leader of BGC-Argo. But it certainly seems possible to get there, he adds. "With a little signal processing, who knows?"

At a minimum, the new method could serve as a backup for BGC-Argo, whose future is in doubt. Several years ago, Honeywell stopped making the pH sensor used by the floats, only to reverse its decision when former President Joe Biden's administration intervened. And although the National Science Foundation funded the first large-scale BGC-Argo deployment, neither the U.S. nor other countries have committed the money needed to keep the program going past next year. At some point, this acoustic method "might be what we've got," Johnson says.

The next step for proving out the technique will be leaving one of these instruments on the sea floor for months at a time, Barclay says. A lower cost version of the instrument might also prove useful for monitoring geoengineering experiments that aim to draw down atmospheric carbon by increasing the alkalinity of coastal waters. It's also possible that the fiber optic cables increasingly being used to sense acoustic signals on the ocean floor could pick up these pH-sensitive frequencies, Lavery says. "If you could set something up for continuous measurements, then it would become pretty useful." □



PHOTO: KATE FLOCK/MASSACHUSETTS GENERAL HOSPITAL

IN OTHER NEWS

Tim Andrews (left) recovered well after receiving a pig kidney earlier this year.

Pig kidney transplant falters

A pig kidney that had been functioning in a New Hampshire man for nearly 9 months has failed, a few days short of setting a record for the longest any organ from another species has worked in a human. In January, surgeons at Massachusetts General Hospital (MGH) gave Tim Andrews a kidney from a pig that had 69 genes altered to make its organs more suitable for use in people. But that editing—and the powerful immunosuppressive drugs Andrews took—apparently weren't enough to stop his body from rejecting the organ. On 23 October, surgeons removed the kidney, which Andrews called "Wilma" because that was the name of his pig donor. Andrews, who is 67, was able to go back on dialysis, but is not eligible for another pig kidney because of the toll the experimental transplant has taken on his health. The xenotransplant record he had been about to pass belonged to Edyth Parker, who in 1964 received a chimpanzee kidney at Charity Hospital in New Orleans and died 1 day before the 9-month anniversary of the transplant. MGH says it is continuing with the experimental transplant of more gene-edited pig kidneys. —Jon Cohen

ARTIFICIAL INTELLIGENCE

At a novel meeting, AI wrote and reviewed the papers

Organizers aim to tune algorithms to reduce peer-review burdens and accelerate science



JEFFREY BRAINARD

Major scientific journals and conferences ban crediting an artificial intelligence (AI) program, such as ChatGPT, as an author or reviewer of a study. Computers can't be held accountable, the thinking goes. But last week, a norm-breaking meeting turned that taboo on its head: All 48 papers presented, covering topics ranging from designer proteins to mental health, were required to list an AI as the lead author and be scrutinized by AIs acting as reviewers.

The virtual meeting, Agents4Science, was billed as the first to explore a theme that only a year ago might have seemed like science fiction: Can AIs take the lead in developing useful hypotheses, designing and running computations to test them, and writing a paper summarizing the results? And can large language models, the type of AI that powers ChatGPT, then effectively vet the work?

"There's still some stigma about using AI, and people are incentivized to hide or to minimize it," a lead conference organizer, AI researcher James Zou of Stanford University, tells *Science*. The conference aimed "to have this study in the open so that we can start to collect real data, to start to answer these important questions." Ultimately, the organizers hope a fuller embrace of AI could accelerate science—and ease the burden on peer reviewers facing a ballooning number of manuscripts submitted to journals and conferences.

But some researchers reject the conference's very premise. "No human should mistake this for scholarship," said Raffaele Ciriello of the University of Sydney, who studies digital innovation, in a statement released by the Science Media Centre before the meeting. "Science is not a factory that converts data into conclusions. It is a collective human enterprise grounded in interpretation, judgment, and critique. Treating research as a mechanistic pipeline … presumes that the process of inquiry is irrelevant so long as the outputs appear statistically valid."

Most of the 315 manuscripts submitted to the conference, which attracted 1800 registrants, were reviewed and scored on a six-point scale by three popular large language models—GPT-5, Gemini 2.5 Pro, and Claude Sonnet 4. The results were averaged for each paper and humans were asked to review 80 papers that passed a threshold score. Organizers accepted 48 covering multiple disciplines, based on both the AI and human reviews.

One accepted paper that organizers highlighted was submitted by Sergey Ovchinnikov, a biologist at the Massachusetts Institute of Technology. His team asked advanced versions of ChatGPT to generate amino acid sequences that code for biologically active proteins with a structure called a four-helix bundle. Scientists typically use specialized software to design proteins. But to Ovchinnikov's surprise, ChatGPT produced sequences without further refinement of his team's query. He and his human colleagues tested two of the sequences in the lab, confirming that a protein derived from one had a four-helix bundle. Still, Ovchinnikov found ChatGPT's performance wasn't flawless. Most of the sequences did not garner "high confidence" on a score predicting whether they would form the desired protein structure.

Data presented at the conference also probed how researchers collaborate with AIs. The organizers asked human authors to report how much work AI contributed in several key areas, including generating hypotheses, analyzing data, and writing the paper. AI did more than half of the hypothesis-generation labor in just 57% of submissions and 52% of accepted papers. But in about 90% of papers, AI played a big role in writing the manuscripts, perhaps because that task is less computationally challenging.

Some human authors who presented at the meeting said their AI partners enabled them to complete in just a few days tasks that usually take much longer, and eased interdisciplinary collaborations with scholars outside their fields. But they also pointed to AI's drawbacks, including a tendency to misinterpret complex methods, write code that humans needed to debug, and fabricate irrelevant or nonexistent references in papers.

Scientists should remain skeptical about applying AI for tasks that require deep, conceptual reasoning and scientific judgment, according to Stanford computational astrophysicist Risa Wechsler, who reviewed some of the submitted papers.

"I'm genuinely excited about the use of AI for research, but I think that this conference usefully also demonstrates a lot of limitations," she said during a panel discussion at the meeting. "I am not at all convinced that AI agents right now have the capability to design robust scientific questions that are actually pushing forward the forefront of the field." One paper she reviewed "may have been technically correct but was neither interesting nor important," she said. "One of the most important things we teach human scientists is how to have good scientific taste. And I don't know how we will teach AI that."

The conference organizers plan to analyze and compare the AI- and human-written reviews. But the remarks garnered by Ovchinnikov's protein-design paper hinted that people and machines may often diverge. An AI reviewer called it "profound." But a human deemed it "an interesting proof-of-concept study with some lingering questions." □

ILLUSTRATION: MINA DE LA O/GETTY IMAGES

eppendorf & Science

PRIZE FOR NEURO BIOLOGY



Meet the Winner 2025

Eppendorf & Science Prize for Neurobiology
Congratulations to Cheng Lyu, Ph.D., Postdoctoral Fellow at Stanford University, School of Medicine, Stanford, USA, on winning the 2025 Eppendorf & Science Prize. Dr. Lyu's winning essay describes the molecular mechanisms how olfactory receptor neurons choose one synaptic partner out of all potential candidates. The olfactory neural circuit of the fruit fly builds its three-dimensional structure by breaking the developmental task into a series of one-dimensional steps. By manipulating this process, Dr. Lyu was able to rewire the neural circuit and altered the animal's courtship behavior. These findings link genetic variation, through circuit assembly, to animal behavior, paving the way to explore how developmental changes influence brain function.

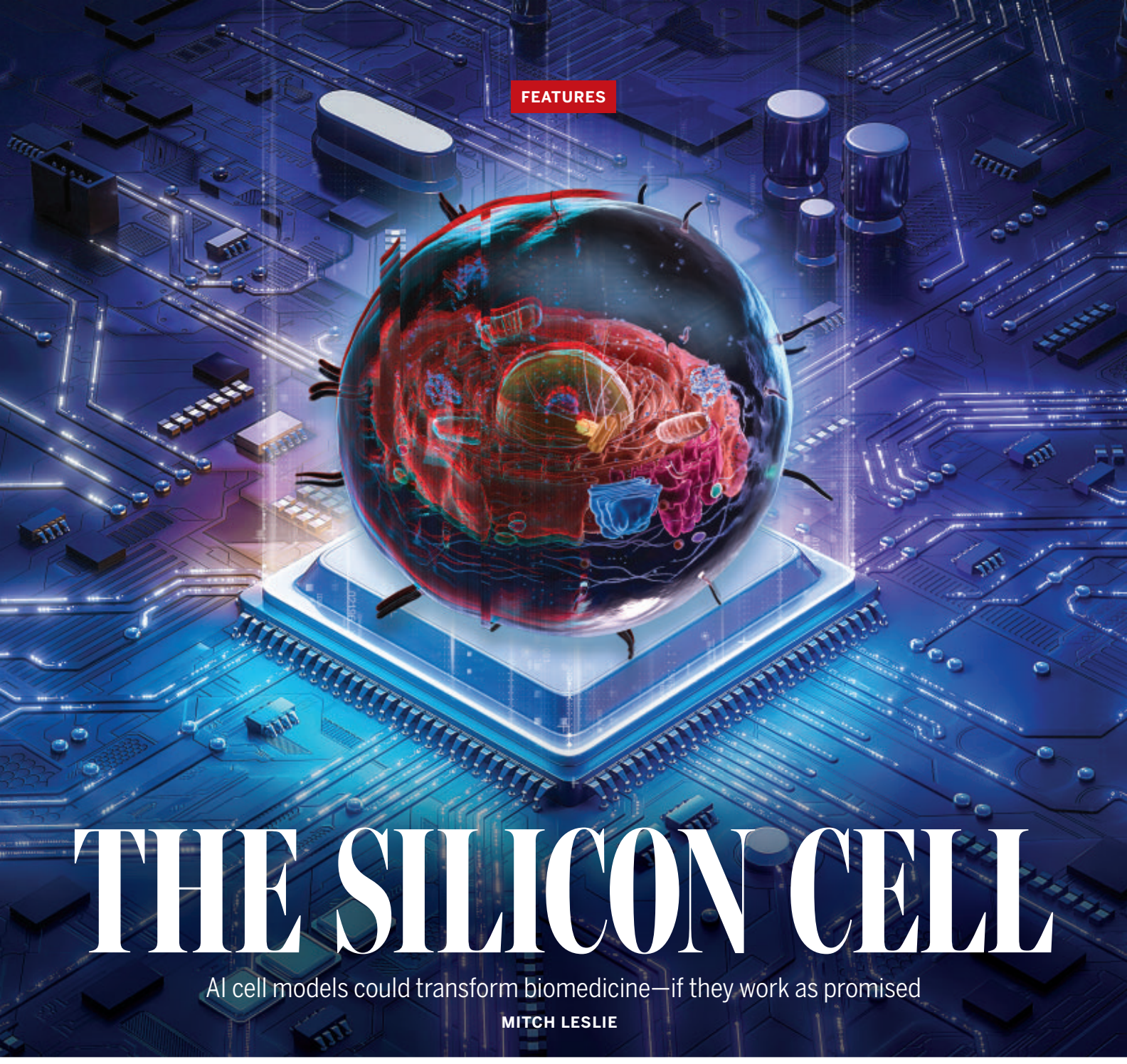
Two finalists are also recognized. Sara Mederos, Ph.D., Wellcome Early Career Fellow at the Sainsbury Wellcome Centre, University College London, for her work revealing how the brain learns to adapt fear responses and suppress them when there is no reason for fear, and Dr. Constanze Depp, Postdoctoral Research Fellow at the Boston Children's Hospital, Harvard Medical School, Broad Institute of MIT and Harvard, for her work examining the role of myelin and oligodendrocytes in Alzheimer's disease.
The annual US$25,000 Eppendorf & Science Prize for Neurobiology honors early career scientists for their outstanding contributions to neurobiological research. The winner and two finalists will be recognized at a Prize Ceremony on November 16, 2025, at the Hard Rock Hotel in San Diego, USA.
If you are interested in joining the Prize Ceremony go to **event.eppendorf.com/prizeceremony-2025** or scan the QR code.

You could be next to win this Prize
If you are 35 years of age or younger and currently performing neurobiological research, you could be next to win this Prize. Deadline for entries is June 15, 2026.

eppendorf Science AAAS

Learn more at: eppendorf.com/prize

AAAS® and Science® are registered trademarks of the American Society for the Advancement of Science, USA. Eppendorf® and the Eppendorf Brand Design are registered trademarks of Eppendorf SE, Germany. All rights reserved, including graphics and images. Copyright © 2025 by Eppendorf SE. Photography: Junjie Zhang.



THE SILICON CELL

AI cell models could transform biomedicine—if they work as promised

MITCH LESLIE

A human cell swarms with trillions of molecules, including some 42 million proteins and a plethora of carbohydrates, lipids, and nucleic acids. Crowded with organelles and other structures, the cell boasts an intricate organization that makes baroque architecture seem plain. Its cytoplasm is a frenzied chemical lab, with molecules continuously reacting, rearranging, and reshaping. In the nucleus, thousands of genes are constantly switching on and off to turn the seeming chaos into concerted actions that help the cell survive and reproduce.

This complexity is more than the human mind can yet fully understand or predict. But many researchers think artificial intelligence (AI), with its prodigious ability to assimilate and process information, might be up to the task. More than 2 decades ago researchers started to build systems of equations meant to simulate some of the cell's workings. Now, they have progressed to AI-driven replicas that, like the large language models taking business and popular culture by storm, ingest vast amounts of data to learn on their own. ChatGPT's attention-grabbing debut nearly 3 years ago inspired the virtual cell builders. "People want this kind of moment for biology," says Kasia Kedzierska, an AI research scientist at the Allen Institute.

IMAGE: 3DARTISTS/SHUTTERSTOCK, ILEXX/ISTOCK. ADAPTED BY N. JESSUP/*SCIENCE*



How soon it is coming depends on whom you ask. Virtual cells that emulate their living counterparts would be a boon for many areas of research. In pharma labs, scientists could use them to quickly evaluate large numbers of potential drugs without the expense and difficulty of experiments. They might serve as test beds for engineering cells to perform novel functions. Virtual cells customized to match a patient's molecular profile could help doctors choose tailored medications. Researchers might even weave cell models into virtual tissues and organs to tackle questions such as how a tumor's environment affects its growth.

Such models could also help researchers make sense of the vast amount of diverse information pouring into molecular databases, says Theofanis Karaletsos, head of AI for science at the Chan Zuckerberg Initiative (CZI). An AI-powered cell mimic, Karaletsos says, "creates an integrated map of knowledge."

Like ChatGPT and its ilk, AI cell models have spawned big promises and hefty expectations. "Whenever a new model appears, it's always the best," says computational biologist Hani Goodarzi of the Arc Institute, who develops such models himself. In June he and more than 20 other researchers launched the Virtual Cell Challenge, a new contest that will put the models to the test annually. Much like a structural biology competition that started in 1994 and helped researchers largely solve the problem of how proteins fold, the Virtual Cell Challenge is meant to spur improvement in a very complex task. For its debut, it is asking AI aficionados to predict the effects of silencing certain genes in human embryonic stem cells.

So far, more than 1000 teams—with names like Cellamander, Zebulon Chow, SmartCell, and Mean Predictors—have entered and are vying for prizes donated by sponsors including Nvidia, the giant tech company that makes the graphics processing units (GPUs) at the heart of many AIs. On 6 December, contest organizers will reveal the final standings, with the top team taking home $100,000 in cash and GPU time. "We want to learn what works and what doesn't work," Goodarzi says.

Even if the models perform well on the test, some scientists expect a long road before they can deliver compelling new science or help biologists. "Despite the hype, [the models] are underperforming," says Alex Lu of Microsoft Research, who studies how AI can identify patterns in biology data. Some seem to have no more predictive power than simpler simulations. The profusion of models is itself a bad sign, says computational biologist Qin Ma of Ohio State University. "One model should be powerful enough that we wouldn't need to see so many of them."

The creators of the models say success is only a matter of time. "We haven't solved the problem yet, but [the approach] is very promising," says Bo Wang, head of biomedical AI for Xaira Therapeutics, a company that hopes to harness the technology for drug discovery.

INSPIRED BY ADVANCES in computing capabilities, researchers began trying to create virtual cells about 25 years ago. They first used computational methods that rely on large sets of equations to recapitulate metabolism, protein synthesis, DNA duplication, and other cell processes.

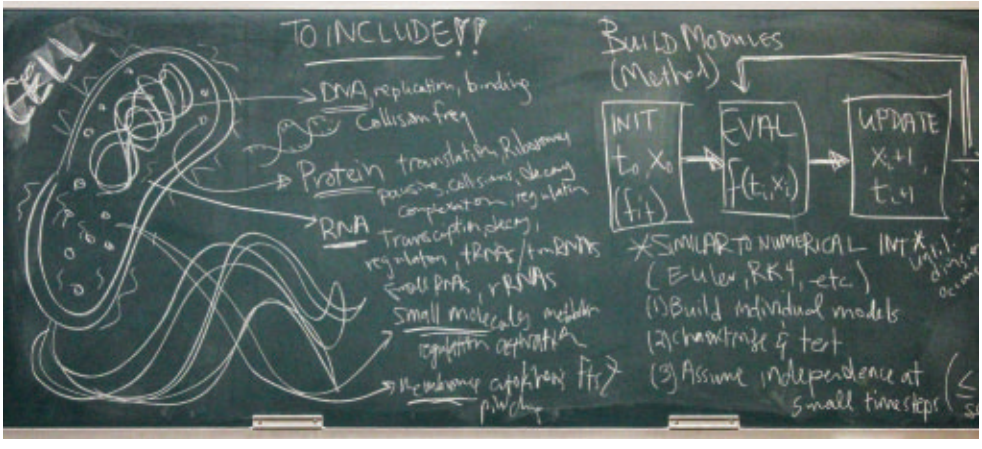
In 2012, Jonathan Karr, now a computational systems biologist at the Mount Sinai School of Medicine, and colleagues in Markus Covert's lab at Stanford University unveiled the first whole-cell model, a silicon version of *Mycoplasma genitalium*. They chose the microbe because it had the smallest genome of any bacterium known at the time: just over 500 genes, compared with more than 4000 in the familiar *Escherichia coli*. To replicate the organism's metabolism, the model calculated the concentrations of more than 700 metabolites as they churn through 1100 chemical reactions. With representations of a chromosome and protein-synthesizing organelles known as ribosomes, the ersatz cell reflected some of the internal structure of its real-life counterpart.

In 2022, cell biologist and biological physicist Zan Luthey-Schulten of the University of Illinois Urbana-Champaign and colleagues unveiled a model that had slightly fewer genes but was more sophisticated. It was based on a minimal bacterial cell, a microbe pared down in the lab to carry a mere 493 genes. But the researchers increased the realism of the model by making it 3D, using data collected by imaging real cells with cryogenic electron tomography. This level of detail came at a computational cost—simulating just 20 minutes of a cell's life required 8 to 10 hours of processing time with high-level GPUs.

Scientists have also developed similar whole-cell models of other bacteria, including *E. coli*. When the

researchers set these models into motion, they can simulate, second by second, a cell's existence as it metabolizes, grows, and divides. The ability to predict short-term cell dynamics, such as changes in concentrations of specific molecules, "is where these models really shine," Karr says. And Luthey-Schulten says researchers will one day be able to create computational models of human cells. "Do I think it will eventually be done? Yes."

Still, Wang says, these non-AI models can only follow biological rules set down by their developers, and it's unclear how closely the mathematical equations that underlie their behavior recapitulate reality. In addition, the models can't handle many of the questions that researchers most want to answer, says computational biologist Fabian Theis of the Technical University of Munich. "It's safe to say they have completely failed" at making certain kinds of predictions, such as forecasting the effects of silencing genes, he says.



This chalkboard from 2012 depicts elements that went into the first published computional model of a whole cell, from Markus Covert's lab.

THE VIRTUAL CELL CHALLENGE reflects hopes that the newer AI-driven models will be more successful. In place of inviolable equations, they learn cells' operating principles on their own, from huge amounts of experimentally derived cell biology data. "We don't teach it any biology," biophysicist Steve Quake of Stanford University says of the AI model TranscriptFormer, which he and CZI colleagues developed.

The models can exploit a torrent of new data, particularly single-cell measurements of gene activity. CZI's CELLxGENE collection, for example, houses gene activity data from more than 35 million human and mouse cells. Model builders also "get to take advantage of a tidal wave of technical development," Karr says.

Like the popular general-use AIs, such as GPT-5, Claude, and Gemini, most AI cell efforts released so far are so-called foundation models, meaning they can in principle probe a wide range of cell processes without further training. So far, AI cell foundation models capture only a sliver of biology and perform a limited set of tasks, such as classifying cells and predicting how they will respond to changes in gene activity. But researchers hope to ultimately combine them into something more comprehensive. "One model that rules them all is the vision," Theis says.

One early cell foundation model was Geneformer, which computational biologist Christina Theodoris of the University of California, San Francisco and her team described in *Nature* in May 2023. The group hoped to improve on a previous, less capable AI model that scrutinized a network of interacting genes and predicted the effects of modifying it. In a 2021 *Science* paper, Theodoris and colleagues reported they had used the model to identify a molecule that may combat a type of heart disease marked by excessive calcium buildup in the organ's valves. Other scientists have been investigating the molecule as a potential therapy. But that success was hard won: The researchers had to deduce the gene network themselves from experimental data and provide it to the model. With Geneformer, "I wanted to see if we could make a foundation model that had a basic knowledge of gene networks," Theodoris says.

She and her colleagues trained Geneformer with measurements of gene activity from nearly 30 million human cells. They found it could infer gene interactions and predict the effects of tampering with specific genes. The model then went through another round of training with gene activity data from heart cells of people who were healthy or who suffered from cardiomyopathy, in which the heart muscle weakens. It then picked out several genes whose inactivation should shift weakened heart cells toward normal operation.

The Geneformer team next went a step beyond most AI research groups and experimentally tested these predictions, using the gene editor CRISPR to disable two of those genes in lab-grown heart muscle cells with mutations that induce cardiomyopathy. Shutting down either of the genes boosted the cells' contraction strength, the researchers found, suggesting drugs that suppress the genes' function could be helpful in people. The results show Geneformer "could point to new directions and accelerate drug discovery," Theodoris says.

Still, Geneformer, like many AI cell models, was trained only on human data, although a few others also incorporate mouse results. TranscriptFormer, which Quake, Karaletsos, and colleagues described in a bioRxiv preprint earlier this year, stands out for its evolutionary breadth. The researchers produced three versions of the model, the most comprehensive of which schooled itself on gene activity data for 112 million cells from 12 species, including humans, mice, nematodes, sponges, and a malaria parasite.

That diverse education may have sharpened its abilities. TranscriptFormer topped scGPT, an AI model Wang and colleagues released in 2024, and Geneformer at classifying rare types of human cells, the researchers reported. The team also found it could classify cells from species that were not in its training set, discriminate lung cells infected by SARS-CoV-2 from healthy cells, and predict the cellular impacts of drugs. These feats are notable, Karaletsos says, because the model received no additional training—it was capable of what researchers call zero-shot learning. TranscriptFormer's

IMAGE: ERIK JACOBSEN/THREESTORY

performance, he says, "shows how people could use tools like this to discover fundamental biology."

Researchers could also integrate cell foundation models into virtual helpers known as AI scientists, which can comb the research literature, devise hypotheses, conduct simulated experiments based on actual data, and write up the results. When computational biologist Marinka Zitnik of Harvard University and colleagues released a set of more than 600 tools earlier this year to help researchers develop AI scientists, TranscriptFormer was among them. Because of the way the model is designed, she says, "results come faster and with higher fidelity."

SO, WHAT'S HOLDING BACK virtual cells? For one thing, most current AI models rely on a single type of data: measurements of gene activity. Researchers are just beginning to include other types of information, such as cell images, that could lead to more powerful and useful models, Ma says. And even though the amount of cell data has exploded in recent years, there's still not enough to fuel the models. "The data limitation is a massive barrier to cell models," Karr says. "There's a limited set of questions we can pursue."

Karr notes that despite the debut of collections like CELLxGENE, many results don't end up in central storehouses where other scientists can easily access them. What the field needs, he says, is a repository like the Protein Data Bank (PDB), which was instrumental in the achievements of AlphaFold, the protein-folding predictor that is probably the most successful AI model in science. (Two of its developers shared a Nobel Prize in Chemistry in 2024.) The PDB houses almost every experimentally determined protein structure in a standard format, in part because most funders and journals required researchers to add any new structure they produced to the database. There's no such requirement for cell data, which are also far more diverse than protein structures, Karr says. "You can't characterize a cell with one technology."

To improve current models, researchers also need a good way to assess how well they work—but scientists are still trying to decide how to do that. After models



> **One model that rules them all is the vision.**
>
> **Fabian Theis**
> Technical University of Munich

complete their initial training, developers often educate them further to improve certain abilities before putting them to the test. That practice mirrors what happens with widely available foundation models, such as GPT-5, which undergo extensive fine-tuning before the public can use them. But some researchers argue that tests should measure a model's zero-shot performance. "If they are inferring a fundamental understanding of biology, it should be true before we fine-tune," Theis says.

A number of studies that gauged the zero-shot performance of recent AI cell models found them wanting. For example, Lu, Kedzierska, and colleagues pitted scGPT and Geneformer against a more primitive AI model and two non-AI methods. To their surprise, the simpler methods outclassed the two foundation models in tasks such as classifying cells, the researchers revealed earlier this year in *Genome Biology*.

Theodoris counters that the tasks had to be relatively easy so the simpler methods could accomplish them. As a result, the tests did not plumb Geneformer's capabilities.

Even with fine-tuning, some AI cell models still fall short, as postdoc Constantin Ahlmann-Eltze of University College London and colleagues reported in *Nature Methods* in August. The challenge they set was to predict how certain genes would respond to alterations in the activity of other genes. They found that a basic mathematical formula provided more accurate results than five cell foundation models and two other AI models. "I was very surprised. Based on the publications, they all looked very good," Ahlmann-Eltze says.

Studies like these test a limited number of models and apply varying criteria. But the field is developing a better sense of what the models should do and how to measure their performance, Kedzierska says. She points to new benchmarking tools that CZI posted earlier this year, which allow researchers to gauge different models' abilities on standard criteria.

The Virtual Cell Challenge that Goodarzi and colleagues have launched has a similar goal. To produce the raw data for the challenge, researchers at the Arc Institute used CRISPR to shut down 300 genes, one at a time, in human embryonic stem cells and then monitored the effects on other genes. Entrants train their models on data for 150 of the genes—and on any other comparable cell data they can find. In the preliminary round, contestants try to predict the effects of silencing a second set of 50 genes. The results go on the contest leaderboard and give the entrants an opportunity to refine their models. The final results will depend on how well the models do on data for a third group of 100 genes. The task is difficult enough that "we don't expect it will be a slam dunk," Goodarzi says.

A few other contests have tried to assess AI models' ability to predict how cells respond to changes, Theis notes. "But it is very good and exciting to see more of these at scale. We need bigger evaluations."

When the contest's results are released in December, they will provide more information to help researchers boost the performance of cell foundation models. And even scientists who have critiqued current virtual cells say they are optimistic about the future of the AI approach. "I wouldn't give up on these models," Kedzierska says. "Where they fail gives us a lot of information." In the end, "I think we will get really good models."

Karaletsos agrees. TranscriptFormer is the "dumbest model we will ever build." □

IMAGE: 3DARTISTS/SHUTTERSTOCK



COMMENTARY

PERSPECTIVES

GENOMICS

How common cuckoos adapt to multiple hosts

Genomic data reveal the complexity of egg mimicry evolution in cuckoos

Michael D. Sorenson[1] and **Claire N. Spottiswoode**[2,3]

Like other obligate brood parasites, common cuckoos (*Cuculus canorus*) reproduce by depositing eggs in "host" nests, thereby exploiting the parental care of another species. After hatching, the cuckoo chick eliminates its competition by ejecting host eggs (and/or chicks) from the nest. Given these extreme costs, hosts evolved the ability to recognize and reject foreign eggs, and cuckoos evolved eggs that mimic the appearance of host eggs to better escape detection (*1*). This coevolutionary arms race is complicated, however, by the cuckoo's exploitation of multiple host species, which presents a puzzle: How does a single cuckoo species evolve eggs that mimic those of many different hosts? On page 527 of this issue, Merondun *et al.* (*2*) report genomic data that reveal a complex geographic mosaic of cuckoo-host coevolution, in which maternally inherited genes, autosomal (biparental) genes, and population structure all contribute to the evolution of a diversity of cuckoo egg phenotypes.

Nearly a century ago, Reginald Punnett (*3*) suggested that maternal inheritance of egg traits through the female-limited W chromosome (female birds have Z and W sex chromosomes, whereas males are ZZ) would allow different female lineages or matrilines to specialize on different host species and evolve distinct, mimetic egg phenotypes, even if males and females mated randomly within a single interbreeding population. Molecular genetic data provide support for this long-standing hypothesis in three different brood parasites. The African cuckoo finch (*Anomalospiza imberbis*) most closely adheres to Punnett's hypothesis, with clear evidence of multiple, host-specific matrilines with different mimetic egg phenotypes (*4*). In the greater honeyguide (*Indicator indicator*), another African species, two highly divergent matrilines are associated with tree- and ground-nesting hosts, respectively, but it is not clear that maternal inheritance explains additional phenotypic variation within each of these major lineages (*5*). Similarly, an earlier study showed that common cuckoo females in Europe and Asia that lay immaculate blue eggs belong to the same ancient matriline (*6*) but left unexplained the genetic basis of the many other egg phenotypes in this species (see the photo).

When avian brood parasites experience diversifying selection associated with their exploitation of multiple hosts, matrilineal diversification and the evolution of new species (speciation) are alternative outcomes that allow for host-specific adaptation (*7*). In both scenarios, the young female brood parasites are assumed to imprint behaviorally on characteristics of their host species (e.g., preferred habitat and/or songs) (*8*, *9*), such that daughters generally parasitize the same species as their mothers. However, speciation requires a mechanism for assortative mating (i.e., mating between male and female parasites reared by the same host species)—which could be enabled by temporal, ecological, and/or geographic segregation of hosts—or host imprinting in both male and female parasites (*8*, *10*). A second critical distinction is that maternal inheritance is effective only for female-limited phenotypes such as egg coloration, whereas speciation allows for host-specific diversification in traits shared by both sexes (e.g., mimetic appearance and behaviors of parasitic chicks) (*11*, *12*). Accordingly, the available data suggest that matrilineal diversification is associated with selection for egg mimicry (*4*–*6*), whereas speciation is associated with chick mimicry (*11*, *13*).

Merondun *et al.* analyzed whole-genome sequencing data for an impressive sampling of cuckoos from across Eurasia to provide improved resolution of matrilineal history and population structure, along with

PHOTO: SWETLANA MESHCHERYAGINA0



The common cuckoo (*Cuculus canorus*) has evolved egg features (top row) to mimic eggs of different host species (bottom row).

the power to identify autosomal loci associated with egg phenotypes. Their analysis confirms prior results that support maternal inheritance of immaculate blue eggs while also identifying a more recent, independent origin of blue-green background color in a matriline mimicking the brambling (*Fringilla montifringilla*), which has blue eggs that are variably patterned with brown markings. Importantly, Merondun *et al.* show that autosomal genes and geography also contribute to variation in egg phenotypes in both the common cuckoo and the closely related oriental cuckoo (*Cuculus optatus*). This suggests that spatial segregation of hosts, at least in portions of each parasite's range, may be sufficient to allow local adaptation based on genes with biparental inheritance. At the same time, and despite the common cuckoo's vast geographic range and array of host species, maternal inheritance of some egg phenotypes may impede speciation by reducing the potential advantages of assortative mating. Overall, the results suggest a geographic mosaic of host-specific adaptation across Eurasia based on multiple mechanisms, rather than a simple model of purely matrilineal diversification. More extensive geographic sampling for other brood parasites will likely reveal similarly complex mosaics (*11*).

For species in which natural selection has favored maternal inheritance as a "solution" for evolving mimicry of multiple hosts, the lack of recombination (i.e., the exchange of genetic material between paired chromosomes during meiosis) in maternally inherited genome components may limit future evolutionary potential, including coevolutionary responses to hosts. In common cuckoos, for example, an adaptive, female-limited plumage polymorphism is also maternally inherited (*14*), but having originated in one matriline, there is no opportunity to recombine this trait with a maternally inherited egg phenotype from another existing matriline. Similarly, maternal inheritance in cuckoo finches, in which different matrilines lay eggs with blue and red background color, respectively, may explain their apparent failure to evolve olive-colored eggs (which require expression of both blue and red pigments), a frequent phenotype in one of their primary host species (*4*).

The lack of recombination also makes it harder to identify the specific genes that underlie differences in egg phenotype. Although Punnett highlighted the female-limited W chromosome, changes in mitochondrial DNA (mtDNA) could also be involved in maternally inherited phenotypes. The W chromosome (except for its pseudoautosomal region) and mtDNA are both nonrecombining and effectively linked by their shared matrilineal history. Thus, the mutations that distinguish divergent matrilines are all correlated with each other, making a standard genome-wide association study ineffective for pinpointing which mutational differences are functionally important. Nonetheless, the suggestion that mitochondria are involved in egg coloration through the heme pathway is an intriguing possibility. Protoporphyrin IX, which is typically responsible for red and brown colors in avian eggshells, is the immediate precursor to heme, which is a component of respiratory complexes II, III, and IV and helps regulate mitochondrial activity, whereas biliverdin, typically responsible for blue or green eggshells, is an intermediate product in heme degradation. Interestingly, one of the autosomal genes that Merondun *et al.* identified as potentially contributing to egg phenotypes, perhaps through interaction with maternally inherited genes, is biliverdin reductase A (*BLVRA*), which converts biliverdin to bilirubin. A presently unknown *BLVRA* allele or nearby regulatory change altering its expression might increase the concentration of biliverdin during eggshell formation. Note that mitochondrial involvement in egg coloration may or may not depend on mutations in mitochondrial genes, as many nuclear genes, including some W-linked genes, are involved in mitochondrial function.

Broad geographic sampling and genomic data allowed Merondun *et al.* to reveal the complexity of host-parasite coevolution in the most famous and well-studied brood parasite. Notably, their results suggest potential interaction between maternally inherited and autosomal genes and, more generally, show that maternal inheritance alone does not account for the diversity of cuckoo egg phenotypes. Learning more about the molecular basis of maternally inherited traits in birds will continue to be challenging. Comparing gene expression in the shell gland, where the eggshell forms and its pigments are deposited, could be informative for egg phenotypes but is infeasible for most wild birds. Comparative genomic analyses that include more brood parasites, their hosts, and other nonparasitic species may do more to expand understanding of host-specific adaptation. Achieving this will require fieldwork on more of the world's ~100 brood parasites, including many species in Africa and Asia that have never been studied intensively in the field. □

REFERENCES AND NOTES

1. N. B. Davies, *Cuckoos, Cowbirds and Other Cheats* (T & A. D. Poyser, 2000).
2. J. Merondun *et al.*, *Science* **390**, 527 (2025).
3. R. C. Punnett, *Nature* **132**, 892 (1933).
4. C. N. Spottiswoode *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2121752119 (2022).
5. C. N. Spottiswoode, K. F. Stryjewski, S. Quader, J. F. R. Colebrook-Robjent, M. D. Sorenson, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17738 (2011).
6. F. Fossøy *et al.*, *Nat. Commun.* **7**, 10272 (2016).
7. G. Dillenseger, *J. Avian Biol.* **2024**, e03252 (2024).
8. R. B. Payne, L. L. Payne, J. L. Woods, M. D. Sorenson, *Anim. Behav.* **59**, 69 (2000).
9. C. Yang, W. Liang, A. P. Møller, *Behav. Ecol. Sociobiol.* **72**, 126 (2018).
10. R. B. Payne, L. L. Payne, J. L. Woods, *Anim. Behav.* **55**, 1537 (1998).
11. N. E. Langmore *et al.*, *Science* **384**, 1030 (2024).
12. G. A. Jamie *et al.*, *Evolution* **74**, 2526 (2020).
13. M. D. Sorenson, K. M. Sefc, R. B. Payne, *Nature* **424**, 928 (2003).
14. J. Merondun *et al.*, *Sci. Adv.* **10**, eadl5255 (2024).

ACKNOWLEDGMENTS

The authors thank C. N. Balakrishnan for comments.

10.1126/science.aec1973

[1]Department of Biology, Boston University, Boston, MA, USA. [2]FitzPatrick Institute of African Ornithology, Department of Biological Sciences, University of Cape Town, Cape Town, South Africa. [3]Department of Zoology, University of Cambridge, Cambridge, UK. Email: msoren@bu.ed

EPIDEMIOLOGY

Broad benefits of the COVID-19 pandemic response

Nonpharmaceutical interventions protected the US population from other infectious diseases

David N. Durrheim

The COVID-19 pandemic exacted an enormous toll in lives and livelihoods, and post-COVID syndrome now afflicts millions of individuals worldwide (*1*). While awaiting the development and rollout of effective vaccines, many countries used a combination of nonpharmaceutical interventions (NPIs), such as restrictions on mass gatherings, school and workplace closures, stay-at-home orders, contact tracing and testing, and requirements for face coverings, to control the spread of the disease and minimize COVID-related deaths. Weighing the costs and benefits of these interventions on physical, mental, and economic health is essential to understand their impact. On page 510 of this issue, Brett and Rohani (*2*) report the effects of NPIs on other common legally notifiable infectious diseases in the US during the COVID-19 pandemic.

Many countries and states underestimated the COVID-19 pandemic threat at a point when aggressive domestic suppression of disease transmission using NPIs, and rapid closure of international borders, could have minimized deaths and allowed a return to near-normal functioning of their economies. A review of Organisation for Economic Co-operation and Development (OECD) country responses to the COVID-19 pandemic demonstrates that countries that responded early with robust NPIs to suppress severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) viral transmission, including Australia, Japan, New Zealand, Norway, Singapore, South Korea, and Taiwan, best balanced death prevention and economic metrics (*3*). However, NPIs had a negative impact on the education and development of young children, and the well-being and mental health of families, particularly in the lowest socioeconomic strata. Concerns have been raised that NPIs implemented during the COVID-19 pandemic could also have disrupted the seasonal and longer-term exposure patterns of other infectious diseases. These disruptions could have led to increased susceptibility of populations to these diseases, resulting in disease resurgence after relaxation of NPIs (*4*).

Brett and Rohani focused on the effects of NPIs on the incidence (rate of disease occurrence in a specified population in a defined period) and trends of other infectious diseases in the US from 2014 to 2024. All diseases included were notifiable by law and were defined as "not rare" (exceeding 1000 notified cases in the US from 2014 to 2022), with no full years of missing data. Animal rabies was the only nonhuman disease included. Compared with other countries, the US has good universal access to diagnostic services and robust disease notification systems, which enabled reliable disease trend analysis. The variable intensity and duration of NPIs across different US states supported thorough comparison of disease trends across the span of the pandemic.

The behavior of the diseases in this analysis statistically coalesced into three distinct clusters that corresponded to their predominant transmission mechanism: environmentally transmitted and zoonotic infections (which spread from animals to humans), sexually transmitted infections, and airborne or aerosol-transmitted infections. Analysis of environmentally transmitted and zoonotic infections, which were not expected to be affected by NPIs, confirmed that surveillance systems for notifiable diseases continued to function adequately during the pandemic. The analysis of sexually transmitted infections demonstrated that NPIs during the pandemic had only a small and nonlasting effect on these infections. Animal rabies and hepatitis A were outliers in the analysis, as their incidence decreased after the pandemic. This trend is explained by a postpandemic wildlife rabies vaccination campaign and a large multistate outbreak of hepatitis A during the 3 years preceding the pandemic.

NPIs resulted in a sharp reduction in the incidence of airborne and aerosol-transmitted infections. Once NPIs were relaxed, their incidence remained lower than that predicted by mathematical modeling in most US states, despite some rebounds. This observation bolsters the "collateral benefit" case for NPIs during pandemic responses. However, the observations and conclusions of the study are US specific, and it would be important to validate them in other countries. As the analysis is based on disease notification data that require well-developed laboratory and diagnostic capacity, this validation would unfortunately be mostly restricted to other high-income settings. Countries across the full spectrum of NPI implementation during the COVID-19 pandemic, including minimal mitigation (e.g., Sweden), partial implementation (e.g., United Kingdom), and elimination (e.g., New Zealand), should be included.

Brett and Rohani did not include measles in their analysis because this disease was rare in the US, where endemic measles was declared as eliminated 25 years ago (*5*). However, measles has resurged in the US and large outbreaks of measles have also been reported in many other countries, particularly in 2024 to 2025. This resurgence suggests that NPIs during the pandemic may have concealed an accumulation of unvaccinated susceptible children before and during the pandemic, which then became apparent once these measures were relaxed. These observations also highlight the fragility of individual countries' measles elimination efforts and the need to set a global eradication goal for this disease (*6*).

COVID-19 vaccines were developed in record time, taking just 326 days from the release of the virus's genetic sequence to market authorization of a safe, effective vaccine. The Coalition for Epidemic Preparedness Innovations, an international organization that funds vaccine development, plans to abbreviate vaccine development against the next pandemic threat to 100 days (*7*). Setting NPIs at a maximum level could be the best approach to smother transmission of the infectious agent while buying time for vaccine rollout. Frameworks for deciding whether a public-health disease elimination approach is warranted have been developed (*8*).

Comprehensive analysis of the full array of positive and negative effects of pandemic control measures enables the identification of the most effective, efficient, and proportionate measures and their optimal timing, targeting, and intensity. This knowledge enables evidence-based guidance for the inevitable next pandemic. □

REFERENCES AND NOTES

1. E. Pazukhina *et al.*, *BMJ Glob. Health* **9**, e015245 (2024).
2. T. S. Brett, P. Rohani, *Science* **390**, 510 (2025).
3. H. Masuhara, K. Hosoya, *J. Policy Model.* 10.1016/j.jpolmod.2025.05.008 (2025).
4. R. E. Baker *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30547 (2020).
5. A. D. Mathis *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **74**, 232 (2025).
6. D. N. Durrheim *et al.*, *Nature* **641**, 1101 (2025).
7. Coalition for Epidemic Preparedness Innovations, "Delivering pandemic vaccines in 100 days: What will it take?" (2022); CEPI-100-Days-Report-Digital-Version_29-11-22.pdf.
8. M. G. Baker, D. Durrheim, L. Y. Hsu, N. Wilson, *Lancet* **401**, 265 (2023).

10.1126/science.aeb5722

University of Newcastle, Newcastle, NSW, Australia. Email: david.durrheim@newcastle.edu.au

Two diverging paths for clean fuel

Fischer-Tropsch synthesis can produce liquid fuels from nonpetroleum sources

Mark Saeys[1,2]

Discovered a hundred years ago, the Fischer-Tropsch synthesis (*1*) has recently been considered as a possible route to produce liquid fuels and chemical building blocks without the use of fossil sources. This chemical reaction converts syngas—a mixture of carbon monoxide and hydrogen—to commercially important chemicals, such as light olefins (*2*) and long-chain hydrocarbons, through a surface polymerization reaction on cobalt or iron catalysts (*3*). However, hydrogen can form water as a by-product, and carbon monoxide can lead to carbon dioxide production. Indeed, the Fischer-Tropsch synthesis produces more unwanted derivatives than hydrocarbons. On pages 516 and 494 of this issue, Cai *et al.* (*4*) and Gao *et al.* (*5*), respectively, report strategies to maximize the incorporation of hydrogen and carbon atoms into hydrocarbon compounds, offering sustainable routes to liquid fuel production with improved efficiency.

In the Fischer-Tropsch synthesis, the removed oxygen from carbon monoxide can lead to two different by-products (see the figure). It first forms water by combining with two hydrogen atoms (*3*, *6*). However, the water molecule can further react with carbon monoxide to produce hydrogen through a water-gas shift reaction that generates carbon dioxide. This reversible side reaction can regenerate hydrogen if coupled to an active catalyst. This even allows hydrocarbon production without hydrogen input (*7*, *8*). Conversely, the reverse water-gas shift reaction permits the Fischer-Tropsch synthesis without carbon monoxide input by using hydrogen and carbon dioxide as the reactants (*9*). Controlling these reactions determines the incorporation of hydrogen and carbon into valuable hydrocarbons.

Catalyst choice is important in determining the activity of the Fischer-Tropsch and the water-gas shift reactions. Cobalt catalysts limit the water-gas shift side reaction and carbon dioxide production (*10*). By contrast, iron catalysts have higher activity toward the water-gas shift reaction than cobalt, which makes them more favorable for converting hydrogen-poor syngas mixtures (with a hydrogen-to–carbon monoxide ratio below 2) to hydrocarbons. However, iron catalysts also reduce the incorporation of carbon atoms into hydrocarbon compounds, instead generating more carbon dioxide. Although the reaction mechanism of cobalt catalysts is relatively well understood (*3*), the iron-based Fischer-Tropsch synthesis remains unclear. Catalytically active sites that are responsible for the water-gas shift reaction are still debated (*11*).

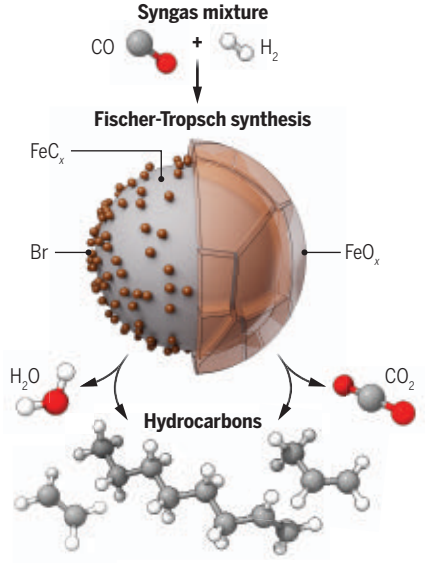
Cai *et al.* showed that feeding trace amounts (<20 parts per million) of bromomethane gas with a syngas mixture reduces water-gas shift activity of iron-carbide catalysts and maximizes carbon incorporation in the hydrocarbon compounds. Bromine atoms covered about 30% of the catalyst surface and reduced the rate of carbon dioxide formation by nearly two orders of magnitude, decreasing carbon dioxide selectivity more than 30-fold. As a result, oxygen atoms were primarily removed as water molecules. Bromomethane also improved selectivity toward the production of olefins, which are important building blocks for the chemical industry. The method of Cai *et al.* improves the carbon atom economy (fraction of carbon from carbon monoxide incorporated in hydrocarbon products) of the Fischer-Tropsch synthesis.

Gao *et al.* coated an iron carbide particle with a thin layer of iron oxide to create a core-shell catalyst. The iron oxide shell maximized the extent of the water-gas shift, which regenerated hydrogen by using the water by-product of the Fischer-Tropsch reaction. This reduced the need for hydrogen in a syngas feedstock, promoting a better hydrogen atom economy (fraction of hydrogen incorporated in the hydrocarbon products). The authors could direct 80% of hydrogen in syngas to olefins. In the method of Gao *et al.*, oxygen atoms from carbon monoxide were mostly removed as carbon dioxide.

Renewable hydrogen production is costly and energy intensive. Driven by the desire to use a range of different carbon sources, such as natural gas, coal, biomass, waste plastic, and carbon dioxide, the Fischer-Tropsch synthesis is experiencing a surge of interest and investment globally. For example, China has commissioned iron-based coal-to-liquid plants to handle a capacity of megatons per year. The electro-sustainable aviation fuel mandates of Europe have driven research in the carbon dioxide–based Fischer-Tropsch synthesis to produce synthetic jet fuel (*12*). The studies of Cai *et al.* and Gao *et al.* demonstrate the rich tunability of iron carbide as a catalyst for the Fischer-Tropsch synthesis. □

Generating sustainable liquid fuels

The Fischer-Tropsch synthesis converts the syngas mixture into liquid hydrocarbons. Introducing a trace amount of bromomethane gas can reduce the amount of carbon dioxide by-product. Alternatively, an iron oxide coating can decrease the formation of water, which improves hydrogen efficiency.



CO, carbon monoxide; CO_2, carbon dioxide; FeC_x, iron carbide; FeO_x, iron oxide; H_2, hydrogen gas; H_2O, water; *x*, number of atoms.

GRAPHIC: N. BURGESS/*SCIENCE*

REFERENCES AND NOTES

1. F. Fischer, H. Tropsch, German Patent DRP 484337 (1926).
2. H. M. Torres Galvis *et al.*, *Science* **335**, 835 (2012).
3. K. T. Rommens, M. Saeys, *Chem. Rev.* **123**, 5798 (2023).
4. Y. Cai *et al.*, *Science* **390**, 516 (2025).
5. C. Gao *et al.*, *Science* **390**, eaea0774 (2025).
6. P. Wang *et al.*, *Nature* **635**, 102 (2024).
7. H. Kölbel, F. Engelhardt, *Erdöl Kohle* **2**, 52 (1949).
8. Y. Xu *et al.*, *Nat. Commun.* **10**, 1389 (2019).
9. B. Yao *et al.*, *Nat. Commun.* **11**, 6395 (2020).
10. B. C. A. de Jong *et al.*, *ACS Catal.* **15**, 10946 (2025).
11. N. S. Featherstone, E. van Steen, *Catal. Today* **452**, 115240 (2025).
12. CARE-o-SENE, Research for a green future: CARE-O-SENE – Catalyst Research for Sustainable Kerosene; https://care-o-sene.com/en/.

ACKNOWLEDGMENTS

The author acknowledges support from the Research Foundation – Flanders (S005226N).

10.1126/science.aec1997

[1]Laboratory for Chemical Technology, Ghent University, Ghent, Belgium. [2]Centre for Advanced Process Technology for Urban REsource recovery (CAPTURE), Zwijnaarde, Belgium. Email: mark.saeys@ugent.be



Chimpanzees living at the Ngamba Island Chimpanzee Sanctuary in Uganda were rescued from the illegal bushmeat trade.

PHOTO: INNOCENT AMPEIRE

PSYCHOLOGICAL SCIENCE

Chimpanzees are natural scientists

Humans and chimpanzees share the potential to rationally revise their beliefs

Brian Hare[1,2,3]

The ability of humans to identify relevant information as evidence and update their beliefs in the face of contradictory evidence is foundational to all scientific inquiry (*1*). The extent to which nonhuman animals are capable of similar forms of reasoning has long been under debate. In his book *The Descent of Man*, Charles Darwin proposed that the greatest challenge for evolutionary theory is human intelligence (*2*). He predicted that psychological elements would be found in nonhumans, particularly in nonhuman apes, that bridge the gap between human and nonhuman cognition (*3*). On page 521 of this

issue, Schleihauf *et al*. (*4*) report that chimpanzees can update their beliefs on the basis of the quantity and quality of new evidence, providing more support for Darwin's radical idea.

An early experimental advance in studying self-knowledge in nonhuman apes was made by hiding a food treat in one of many hollow tubes and observing when the animals peeked inside it. Although they rarely peeked if they had watched the baiting, they frequently visually inspected the tubes when they had not observed it (*5*). Moreover, nonhuman apes were more likely to check the tubes when the stakes were higher (with more desirable food hidden) or when a delay that increased their likelihood of forgetting was introduced before they searched (*6*). These experiments also demonstrated that peeking does not come from practice. When faced with uncertainty, the majority of juvenile and adult chimpanzees spontaneously sought information in this way on their first exposure to the experimental setup (*7*, *8*). Great apes (human and nonhuman) have also shown some capacity to revise a choice when confronted with contradictory evidence. In a different set of experiments, chimpanzees made an initial selection between two boxes containing food of varying desirability. A magnifying glass was then introduced, altering the appearance of the food. When the new evidence was inconsistent with prior information, chimpanzees often peeked into the boxes from another angle before making a final decision, which suggests that they might have revised their beliefs when confronted with conflicting evidence (*9*).

...chimpanzees consistently made rational choices two or three times more often than nonrational ones in all scenarios.

Despite this progress, it has remained unclear whether chimpanzees possess the capacity to rationally evaluate the strength of new evidence in relation to their existing beliefs and revise those beliefs only when appropriate, as humans do. This capacity for reflective responsiveness to reasons is likely to be the only definition of rationality that the strictest opponents of animal rationality will admit (*1*). Schleihauf *et al*. adopted a previously developed experimental setup in which subjects are given a chance to revisit an initial choice after being presented with additional evidence (*9*). Across five experiments, 15 to 23 chimpanzees living at the Ngamba Island Chimpanzee Sanctuary in Uganda were tested on their ability to revise decisions when presented with varying strengths of evidence regarding the location of hidden food rewards.

Schleihauf *et al*. gave subjects eight opportunities to locate food hidden in two identical containers. Chimpanzees could select one of the containers but not both. For each trial, they received two sequential clues: one providing strong, conclusive evidence, and the other offering weak, inconclusive evidence. In half of the trials, the strong clue was presented first, followed by the weak; in the other half, the order was reversed. Chimpanzees consistently based their choices on the strong form of evidence, regardless of the order or modality (visual, when food could be seen inside the container, or auditory, when the container made a noise indicative of food when shaken) in which they received the information. Notably, when weak evidence was encountered first, subjects frequently revised their choice after being exposed to the stronger evidence, which is consistent with rational revision of beliefs.

To determine whether chimpanzees were simply discarding weaker evidence in favor of stronger evidence, three containers were used: One container provided auditory evidence, another provided visual evidence, and the third provided no evidence. Critically, the container with the strongest visual evidence of food was removed after presentation. If subjects had simply discarded weaker evidence once encountering stronger evidence, then their choices of a container should have been random. Instead, chimpanzees chose the container with weaker auditory evidence, suggesting that they integrated the relative strength of all evidence rather than overwriting weaker clues.

The study also revealed that chimpanzees are sensitive not only to the type but also to the quantity of evidence available to support their choices. Subjects were first exposed to weak evidence (a sound resembling food being dropped) and then stronger evidence (a treat visible in another container). As expected, subjects initially chose the container with the visible treat. However, when provided with additional auditory evidence suggesting that two treats might be hidden together, they revised their decision and switched away from the container where they had seen only one treat. This change occurred more often than in a control condition in which subjects only received auditory evidence that reaffirmed that the food they initially heard was still present in the container.

What about the effect of misleading evidence? Subjects were shown food in a container through a transparent cover only to discover that what they had initially seen was not a treat but a picture of a treat painted on the cover. Chimpanzees switched away from the misleading evidence more often than in a control condition in which the transparent cover was unmarked.

In the experiments by Schleihauf *et al*., chimpanzees consistently made rational choices two or three times more often than nonrational ones in all scenarios. These findings demonstrate that nonhuman apes have a spontaneous ability to weigh evidence of varying strengths, revise prior choices, and adapt when evidence is revealed to be unreliable.

Science fiction author Arthur C. Clarke allegedly stated that "either we are alone in the Universe or we are not. Both are equally terrifying." The finding that other great apes are capable of self-reflection means that humans are not alone as rational beings. How tragic it would be if, just as humans discover this shared rationality, they bring about the loneliness that Clarke may have found so terrifying. Humanity should not ignore, so soon after primatologist Jane Goodall's passing, that the chimpanzees in Schleihauf *et al*.'s study were rescued by her and others from the illegal bushmeat trade that threatens their species' survival (*10*). Meanwhile, it is a profound gift to know that when a chimpanzee stares into a person's eyes, they too could be reflecting on their beliefs about humans. □

REFERENCES AND NOTES

1. G. Melis, S. Monsó, *Philos. Q.* **74**, 844 (2024).
2. C. Darwin, *The Descent of Man*, vol. 1 (J. Murray, 1871).
3. B. Hare, "The Darwinian road to morality" in *A Most Interesting Problem*, J. Desilva, Ed. (Princeton Univ. Press, 2021).
4. H. Schleihauf *et al*., *Science* **390**, 521 (2025).
5. J. Call, M. Carpenter, *Anim. Cogn.* **3**, 207 (2001).
6. J. Call, *Anim. Cogn.* **13**, 689 (2010).
7. A. G. Rosati, B. Hare, *Biol. Lett.* **7**, 15 (2011).
8. A. G. Rosati, E. Felsche, M. F. Cole, R. Atencia, J. Rukundo, *Cognition* **251**, 105898 (2024).
9. C. O'Madagain *et al*., *Proc. Biol. Sci.* **289**, 20212686 (2022).
10. S. R. Ross, J. G. Leinwand, *Biol. Lett.* **16**, 20200033 (2020).

ACKNOWLEDGMENTS

The author acknowledges support from the Templeton World Charity Foundation (funder DOI 501100011730), grant TWCF-2022-30062.

10.1126/science.aeb7565

[1]Department of Evolutionary Anthropology, Duke University, Durham, NC, USA. [2]Department of Psychology & Neuroscience, Duke University, Durham, NC, USA. [3]Center for Cognitive Neuroscience, Duke University, Durham, NC, USA. Email: apeminds@gmail.com

POLICY FORUM

INNOVATION

How collaboration with China can revitalize US automotive innovation

Strategic collaboration, not protectionist barriers, should be pursued

John Paul Helveston

The United States stands at a crossroads in its approach to China's rapidly advancing plug-in electric vehicle (PEV) industry, one that will shape not only domestic economic competition but also global climate goals, the future of automotive innovation, and the principles that will govern international trade in an era of rapid technological transformation. Although current protectionist US policies may temporarily benefit the Big Three US automakers, they can also stifle competitive pressure that drives innovation. The European Union offers an example of how a more collaborative approach can lead to important local investments that accelerate PEV development. Rather than retreating behind protectionist barriers, US policy-makers and industry leaders should pursue strategic collaboration that leverages China's manufacturing scale and cost advantages while preserving US strengths in advanced research.

Ford chief executive officer Jim Farley recently provided stark evidence of the competitive gap with China after personally evaluating a Xiaomi SU7—a Chinese battery electric vehicle (BEV) priced at around $30,000 with a driving range of more than 450 miles [the PEV category includes both BEVs and plug-in hybrid vehicles (PHEVs)] (*1*). The verdict was unambiguous: Chinese manufacturers are now producing superior electric vehicles at considerably lower price points than those of US manufacturers. Market data unequivocally support this conclusion. Whereas Chinese manufacturers offer numerous options under $40,000 with competitive driving ranges of 300 to 450 miles, US offerings remain concentrated in premium segments with considerably higher pricing for comparable ranges (see the figure). Chinese BEV makers are also pushing the frontier of achievable charging speeds and driving ranges on a single charge. Both BYD and CATL (Contemporary Amperex Technology Co., Limited) have produced batteries capable of charging at rates as high as 1 MW, translating to 250 to 325 miles of range in just 5 min. (*2*). The highest-range BEV in China, the Zeekr 001, delivers 645 miles at $56,426, whereas the sole US BEV that exceeds 500 miles—the Lucid Air—costs $169,000 for 517 miles (*3*). In 2024, China had 52 BEV models with ranges greater than 400 miles; the United States had four, all priced above $75,000 (*3*).

With high driving ranges, rapid charging speeds, and low prices, BEVs are quickly becoming the dominant vehicle technology in China, with PEV sales reaching 12.86 million in 2024 (46.7% of new vehicle sales) (see fig. S1) compared with just 1.56 million (9.8%) in the United States (*3*). China has also emerged as a net vehicle exporter, exporting 1.28 million PEVs in 2024—nearly as much as domestic US PEV sales (*3*).

The US auto industry faces a concerning future, with implications that could reverberate throughout the US economy. Protectionist policies may not only stifle domestic innovation, they also cannot prevent losses abroad where Chinese PEV manufacturers are rapidly gaining ground. The trajectory of General Motors (GM) illustrates this challenge: Once a leading manufacturer in China, the company finished 2024 ranked 16th in sales, going from highly profitable (25% margins) to losing money on every sale (–20.5% margins) and acknowledging that its Chinese operations are worth $5 billion less than previously estimated (*4*). The United States risks becoming an "island of tailpipes," isolated from global automotive innovation and dependent on domestic sport utility vehicle and truck sales while forfeiting historical global profits derived from international markets. The employment impacts could be substantial; the auto industry supports an estimated 10.1 million US jobs (about 5% of private-sector employment) and generates $730 billion in annual paychecks (*5*).

The path to revitalizing the American auto industry requires strategic engagement with the forces reshaping the global automotive landscape through carefully structured agreements that address security concerns, protect worker interests, and ensure reciprocity in market access. The success of such engagement depends on the willingness of China to collaborate on more balanced terms than in previous decades and the willingness of the United States to allow innovative Chinese firms market access while managing risks instead of imposing wholesale restrictions.

HOW CHINA GOT AHEAD

China's current leadership in PEVs is the result of a long history of factors, not the least of which is a 20-year policy strategy that included forced technology transfer requirements for foreign automakers, preferential treatment for domestic battery suppliers, and an estimated $230 billion in subsidies between 2009 and 2023 (*6*). Local-level policies were also critical; in many Chinese cities, obtaining a license plate for an internal combustion engine vehicle (ICEV) involves lengthy wait times and high costs, whereas PEV plates are often free and immediately available. PEV drivers also enjoy access to many parts of cities where ICEVs face restrictions. Last, Chinese firms have made substantial investments into production innovations and supply chain development (*7*), scaling up critical material processing and battery manufacturing (*8*) while pioneering automation techniques that have reduced vehicle-development time to 18 months—nearly half that of typical incumbents.

Despite pioneering electric vehicle development with GM's EV1 in 1996 and Tesla's innovations in the 2000s, the United States has struggled to develop a competitive PEV industry to counter Chinese market dominance. Policy inconsistency made it difficult for automakers to justify large investments into building a comprehensive PEV supply chain (*9*). Consumer preferences are also important, with re-

Summary of public charge points in China and the United States as of January 2025

	Total public chargers		Charges per 100 EVs	
	United States	China	United States	China
All charge points	200,000	3,580,000	3.1	10.4
Fast (dc) charge points	50,000	1,600,000	0.8	4.7
Slow (level 2) charge points	150,000	1,980,000	2.3	5.8

DATA: GLOBAL EV OUTLOOK 2025; ELECTRIC VEHICLE CHARGING – INTERNATIONAL ENERGY AGENCY (*4*)

search showing that US consumers are on average less willing than their Chinese counterparts to adopt a PEV (*10*). Electricity rates are also considerably lower relative to gasoline prices in China compared with those in the United States, making the operational cost advantages of PEVs more pronounced in China. The United States also has fewer local-level policy tools to promote PEV adoption beyond preferential parking or access to restricted lanes on highways. Last, the United States has struggled to keep pace with China's charging infrastructure construction. The $5 billion National Electric Vehicle Infrastructure program exemplifies this struggle, with only 19 stations built over 2 years because of coordination challenges between installers, utilities, and other stakeholders (*11*). By the end of 2024, the United States had installed just 200,000 public charging points compared with China's 3.58 million (*2*), translating to more than three times as many chargers per EV in China than in the United States (see the table).

The United States risks becoming an "island of tailpipes"...

COMPETITIVENESS THROUGH COOPERATION

The disparities in PEV market offerings between China and the United States reflect not merely differences in policy priorities but a fundamental divergence in innovation and manufacturing capabilities that raises an urgent question: How can the United States gain access to affordable, high-performance PEVs? The answer may lie not in protectionism but rather in strategic collaboration with Chinese manufacturers.

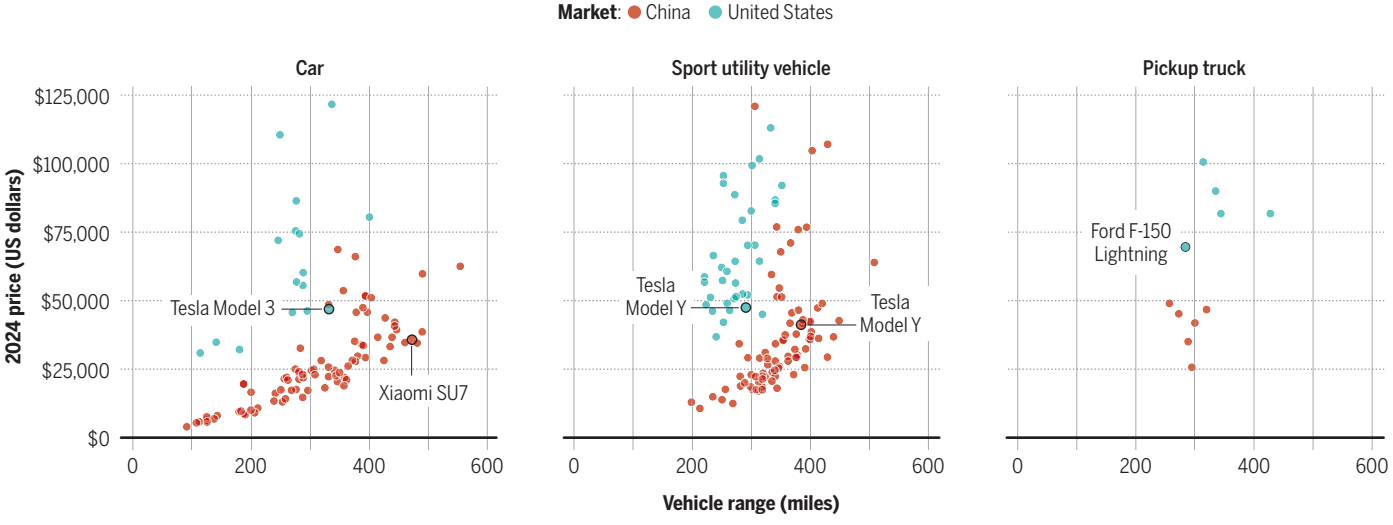
Strategic collaboration can facilitate the transfer of leading battery and PEV technology from China without necessarily creating over-dependency. One path is to pursue technology licensing agreements with Chinese firms that have developed cost-effective battery and vehicle technologies. The discussions between Ford and CATL, for example, on licensing battery technology represent a promising model that would give Ford access to high-performance battery technology while maintaining local manufacturing in the United States. Likewise, the United States could leverage access to its large domestic market as a negotiating tool to secure more favorable terms for US companies operating in China while facilitating Chinese investment in the United States. A strategic integration with Chinese manufacturing expertise through foreign direct investment (FDI) by leading Chinese firms could improve the cost competitiveness of US-made vehicles. Investments by Gotion Battery in Illinois and Michigan demonstrate how Chinese firms can bring capital and expertise to the United States while creating local jobs for US auto workers. Strategic coordination of these investments, rather than blanket restrictions, would better serve long-term US interests.

Furthermore, the United States maintains substantial advantages in fundamental research, software development, and advanced materials, whereas China excels in manufacturing scale, commercialization, and cost reduction (*8*). Fostering formal and informal innovation networks that connect these complementary strengths could accelerate progress in next-generation technologies such as solid-state batteries and autonomous systems more effectively than either nation could achieve independently. For example, the United States and China could establish a dedicated advanced transportation technology dialogue focused on technical standards, regulatory harmonization, and targeted joint research initiatives, potentially leveraging the recently renewed US–China Science and Technology Agreement. This government-to-government engagement would create a supportive framework for private-sector collaboration while addressing legitimate concerns about intellectual property protection and market access.

Equally important is to revise existing policy to facilitate these opportunities. Current Foreign Entity of Concern (FEOC) restrictions—which ban Chinese, Russian, Iranian, and North Korean companies from US EV supply chains and federal incentives—should be refined to allow case-by-case security reviews for manufacturers seeking to establish US facilities. Restoring and expanding the Inflation Reduction Act (IRA) incentives—particularly the $7500 consumer tax credits for EV purchases that have been restricted by FEOC rules—would help to establish a domestic market for more affordable PEVs while preserving the IRA's effectiveness in stimulating domestic investments. The IRA spurred $277 billion in announced private investments for battery and PEV manufacturing facilities across the United States, supporting an estimated 350,000 manufacturing jobs (*12*). Likewise, when evaluating Chinese investments, the Committee on Foreign Investment in the United States review process should distinguish between strategic and nonstra-

Prices versus driving ranges for BEVs available in the United States and China

Chinese manufacturers offer many model year 2024 battery electric vehicles (BEVs) at more affordable prices with higher driving ranges on a single charge. US offerings remain concentrated in premium segments with higher pricing for comparable range performance. Data include BEV specifications collected from autocango.com for China and carsheet.io for the United States (*3*).



GRAPHIC: V. PENNEY/*SCIENCE*; (DATA): AUTOCANGO/AUTOCANGO.COM; CARSHEET/CARSHEET.IO

tegic technologies, such as advanced semiconductors or defense systems versus commercial EV manufacturing. Rather than relying solely on tariffs—which provide temporary market protection but no access to cutting-edge technologies—the United States could implement a calibrated approach combining modest protective tariffs sufficient to encourage foreign investment with streamlined approval processes for companies willing to manufacture domestically and share technological expertise with US partners.

LESSONS FROM SUCCESSFUL COLLABORATIONS

Historical precedent supports a more balanced approach that embraces FDI from Chinese firms to revitalize local manufacturing while creating jobs. When Japanese automakers such as Toyota established US production facilities in southern US states beginning in the 1980s, initial concerns about job displacement gave way to recognition of substantial benefits. US automakers learned from the Toyota Production System, adopting lean manufacturing principles that improved their own capabilities and efficiencies. These investments brought not just jobs but knowledge transfer that strengthened the entire US automotive ecosystem. A similar dynamic could occur with Chinese EV investments, in which US firms could gain exposure to cutting-edge manufacturing techniques and supply chain innovations developed in China.

Tesla's entry into China is another example that demonstrates how engagement can yield mutual benefits rather than zero-sum outcomes. When Tesla established its Shanghai Gigafactory in 2019, local suppliers upgraded their capabilities to meet Tesla's quality requirements, all while reducing manufacturing costs. The Shanghai facility became Tesla's most efficient factory, producing vehicles at substantially lower costs than that of its US operations and improving the company's global competitiveness by applying that manufacturing knowledge to newer facilities. The competitive pressure brought by Tesla benefited both sides: Tesla achieved profitability, and Chinese manufacturers benefitted from better-quality suppliers, sparking an innovation cycle that drove down costs industrywide. The first Tesla BEVs rolled off the Shanghai facility lines in December 2019, and PEV prices in China fell below ICEV prices on average for the first time just 2 years later, leading China's government to phase out consumer purchase subsidies for PEVs (*2*).

Crucially, Tesla's experience shows that security concerns can be managed through targeted solutions rather than blanket restrictions. China was concerned that Tesla's vehicles—equipped with cameras and sensors for navigation—could potentially collect sensitive visual data near government buildings, military installations, or high-technology zones. To address these concerns, Tesla agreed to specific requirements: storing all vehicle data collected in China within Chinese data centers (preventing potential US government access), restricting vehicle operations in designated sensitive areas, and using only Chinese government–approved mapping and navigation systems. Only after implementing these cybersecurity measures did the Chinese government grant Tesla permission to activate its Full Self-Driving technology. By contrast, the United States has used cybersecurity concerns to justify broad restrictions specifically targeting Chinese companies, including legislation that categorically blocks Chinese EV manufacturers from the US market (*13*). Although security risks are real, blanket exclusions are counterproductive. A more effective approach would develop transparent security standards and compliance verification that all connected vehicles must meet, regardless of origin (*14*).

In Europe, we can see promising examples of such cooperation. By opening to investments from Chinese firms, incumbent automakers are gaining access to the latest battery technologies, improving their competitiveness in the European market. CATL is building a $7.6-billion battery plant in Debrecen, Hungary, expected to produce 100 GWh of batteries annually for European automakers, including Mercedes-Benz, BMW, and Volkswagen. Similarly, BYD is constructing its first European electric vehicle manufacturing facility in Hungary to produce both BEVs and PHEVs specifically for the European market.

THE STAKES ARE GLOBAL

An isolationist approach in the United States undermines both climate and innovation objectives. Light-duty vehicles generate 16% of US carbon emissions (*2*), making the rapid deployment of affordable PEVs essential for meeting climate targets. Furthermore, isolating one of the world's largest automotive markets from leading PEV technology could have profound negative consequences for global innovation. Solid-state batteries, advanced materials, and more efficient motors could all develop more rapidly through collaborative networks than through parallel, isolated efforts. Research on wind and solar industries demonstrates that cross-border partnerships accelerated innovation while allowing each country to maintain distinctive industrial specializations when appropriate collaborative frameworks were established (*15*).

At the center of the PEV revolution lies a nexus of advanced computing, data security, and critical minerals that requires engagement to develop sustainable solutions. By developing a policy strategy that fosters FDI while protecting core interests, the United States can accelerate its electric transition and maintain a globally competitive automotive industry. The alternative—excluding Chinese technology and expertise—risks leaving US automakers uncompetitive while delaying the urgent global transition to sustainable transportation. □

REFERENCES AND NOTES

1. E. Halper, *The Washington Post* 25 April 2025.
2. International Energy Agency, "Global EV outlook 2025" (International Energy Agency, 2025); https://www.iea.org/reports/global-ev-outlook-2025/.
3. J. P. Helveston, jhelvy/us-china-pev-policy-2025: Main release. Zenodo (2025); https://zenodo.org/records/16878932.
4. K. Bradsher, *The New York Times* 19 December 2024.
5. Alliance for Automotive Innovation, "State of the US automotive industry" (Industry Report, Alliance for Automotive Innovation, 2025).
6. S. Kennedy, "The Chinese EV dilemma: Subsidized yet striking" (Center for Strategic and International Studies, 2024); https://www.csis.org/blogs/trustee-china-hand/chinese-ev-dilemma-subsidized-yet-striking.
7. J. P. Helveston, Y. Wang, V. J. Karplus, E. R. H. Fuchs, *Res. Policy* **48**, 206 (2019).
8. J. Helveston, J. Nahm, *Science* **366**, 794 (2019).
9. J. D. Graham, K. B. Belton, S. Xia, *Issues Sci. Technol.* **37**, 72 (2021).
10. J. P. Helveston *et al*. Transport. Res. A Pol., *Pract.* **73**, 96 (2015).
11. D. Ferris, *E&E News* 8 October 2024.
12. J. Turner, The Big Green Machine: Tracking the North American Clean Energy Supply Chain (2025); https://www.the-big-green-machine.com.
13. C. D. Miller, End Chinese Dominance of Electric Vehicles in America Act of 2024, H.R.7980 — 118th Congress (2023-2024); https://www.congress.gov/bill/118th-congress/house-bill/7980.
14. M. R. Davidson, V. J. Karplus, J. I. Lewis, J. Nahm, A. Wang, *Science* **377**, 1266 (2022).
15. J. Nahm, *Collaborative Advantage: Forging Green Industries in the New Global Economy* (Oxford Univ. Press, 2021).

ACKNOWLEDGMENTS

The author thanks the Fudan-UC Center on Contemporary China at the University of California, San Diego's School of Global Policy and Strategy for continuing to support a meaningful dialogue between the United States and China, and specifically the workshops on climate and United States–China relations organized by M. Davidson and L. Guang; this paper was greatly informed by conversations at these workshops. The author also thanks E. Wilson, J. Busby, and P. Triolo for their comments and suggestions. Anthropic's Claude 4.0 Sonnet AI model was used for editorial assistance purposes only during the preparation of this manuscript. All AI-generated content was reviewed, edited, and verified for factual accuracy and citation validity by the author. The author takes full responsibility for the final manuscript.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adz0541

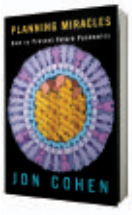
10.1126/science.adz0541

Department of Engineering Management and Systems Engineering, George Washington University, Washington, DC, USA. Email: jph@gwu.edu

PUBLIC HEALTH

More than a miracle

Pandemic preparedness is as essential as ever, argues a journalist **Caitlin Rivers**



Jon Cohen
Knopf, 2025.
448 pp.

PLANNING MIRACLES | *Planning Miracles*, by Jon Cohen—a veteran journalist at *Science*—introduces readers to the world of pandemic preparedness at a chaotic time for the field. The United States is weathering an erosion of public health institutions, with far-reaching rollbacks diminishing funding, legal authorities, and trust. The book reminds readers of the distance humans have traveled to secure good health and illustrates the dangers that may befall us should our commitment to those protections wane. It is both a celebration of scientific achievement and a warning about the costs of complacency.

The book's title comes from one of history's most welcome achievements: the development of the Salk polio vaccine in 1955. March of Dimes founder Basil O'Connor described the vaccine's development as a planned miracle. "None of us dared to believe 17 years ago that this would happen in less than a century," O'Connor wrote that year in the introduction to the foundation's annual report. "And yet when it happened, it took everyone's breath away."

Cohen contrasts these miracles—which is to say, the deliverances of scientific breakthroughs and the scientists who accomplish them—with the tolls exacted by failure. "The opposite of a miracle is a disaster," he writes, "and few disasters exact a higher price than a pandemic."

The book's narrative is built around a theme captured in Benjamin Franklin's famous aphorism "an ounce of prevention is worth a pound of cure," coined after witnessing the abysmal response mounted by a ragtag volunteer firefighting effort to a destructive fire in Philadelphia. Throughout, Cohen successfully argues that investing in pandemic preparedness is as pragmatic and essential as investing in professional firefighting. The book weaves together historical accounts of the characters and breakthroughs that built our current armamentarium with firsthand interviews with contemporary scientists.

"Outbreaks are inevitable," Cohen writes. "If we want to prevent them from becoming epidemics…we must improve our understanding of what sparks them and evolve our responses, too." To that end, the book covers an enormous breadth, managing to survey a myriad of pandemic-prone pathogens, including influenza, the coronaviruses, and filoviruses, and exploring the scientific tools to combat each, in turn. Readers join Cohen as he tours a startup in Boulder, Colorado, developing the next generation of respirators; they tag along as he probes the enormously complicated and, at times, political nature of international clinical trial planning; and they follow him into mountains and rainforests where new viruses are found.

The most timely chapters examine viral threats that continue to knock at humanity's door. Influenza, particularly the avian influenza A(H5N1) that infected millions of birds before expanding into dairy cows, is one such hazard. Cohen traces the long history of influenza preparedness efforts, noting that H5N1 has threatened "ruination" since its emergence in 1996. This history will be of particular interest to readers seeking to go behind the headlines. His treatment of mpox follows a similar arc, showing how the virus transitioned from a rare zoonotic disease to a global outbreak in 2022 that continues even today. Together, these chapters illustrate Cohen's key insight: that the viruses we know about, and have tools to address, still pose serious threats when surveillance, countermeasures, and response capacities are neglected.

The book does not shy away from technical details. Cohen's treatment of the science behind pandemic preparedness is extensive and sophisticated, which may be a challenge for some readers. He delves into structural virology, immunology, genomics, and more, with technical passages sometimes extending for several pages. However, he balances these sections with human stories of the researchers doing the work, which should help sustain momentum for nonexpert readers.

Cohen also departs somewhat from the book's main themes on countermeasures to argue strongly in favor of a natural origin of severe acute respiratory syndrome coronavirus 2. He includes a defense of Peter Daszak—leader of the nonprofit organization EcoHealth Alliance, whose work on coronoviruses with a virology laboratory in Wuhan, China, has been heavily scrutinized as a potential origin of the COVID-19 pandemic. Not all readers will find Daszak to be a sympathetic character. Still, the inclusion of more contentious topics furthers the book's effort to portray science as a purposeful endeavor populated by real, complex people.

Planning Miracles will appeal to readers interested in public health, pandemic preparedness, and the history of medicine—particularly those hungry for scientific detail. The book's central argument about the importance of investing in pandemic preparedness resonates, especially now, as the devastation from the COVID-19 pandemic fades into memory and is replaced by corrosive skepticism and indifference. In reminding us that miracles must be planned, Cohen delivers an essential message for our tumultuous times. □

10.1126/science.aeb9832

The reviewer is director of the Center for Outbreak Response Innovation, a senior scholar at the Center for Health Security, and an associate professor at Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. She is also the author of *Crisis Averted: The Hidden Science of Fighting Outbreaks* (Viking, 2024). Email: crivers6@jhu.edu

PODCAST



A composite image shows a star's death in the constellation Taurus.

IMAGE: NASA/JPL-CALTECH/ESA/CXC/UNIV. OF ARIZ./UNIV. OF SZEGED

SCIENCE OF DEATH

Rounding out our 2025 series on the science of death, this week on the *Science* podcast, astrophysicist Katie Mack, author of **The End of Everything (Astrophysically Speaking)**, succinctly summarizes the ways our Universe may eventually end, placing our existential concerns about the end of life as we know it into cosmological perspective.

https://scim.ag/TheEndOfEverything

10.1126/science.aec7043

PHYSICS

Extraterrestrials illustrated

A physicist and a cartoonist join forces with an accessible guide to thinking about life in the cosmos

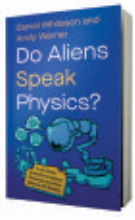
Matt von Hippel



The Drake equation estimates the number of alien civilizations whose messages we should be able to receive.

Since Randall Munroe's *What If?* published in 2014, illustrated pop science books have earned a niche on bookstore shelves. And their subject matter has gradually grown in ambition, as cartoonists and scientists work together to present increasingly well-researched takes on fascinating and obscure topics, with Kelly and Zach Weinersmith's 2023 book, *A City on Mars*, a standout recent example (*1*).

Do Aliens Speak Physics? brings together two veterans of the form: Daniel Whiteson, a physicist who has written two previous books with cartoonist Jorge Cham, and Andy Warner, a cartoonist who has previously covered weighty topics, such as Lebanon's Cedar Revolution (*2*). Like the Weinersmiths, Whiteson and Warner are trying something ambitious and well researched, attempting to convey a topic much deeper than the book's whimsical illustrations of goo-dripping aliens would suggest. To spoil my conclusion: They don't quite get there, but the questions they ask may justify reading the book, regardless.



Do Aliens Speak Physics?
Daniel Whiteson and Andy Warner
Norton, 2025. 272 pp.

The questions posed by Whiteson and Warner stem from their book's deceptively simple title: Do aliens speak physics? Historically, many physicists assumed so. When Frank Drake and Carl Sagan composed missives meant for aliens, such as the Pioneer plaque (*3*) and the Arecibo message (*4*), they began with references to physics and chemistry, invoking ideas such as hyperfine transitions in hydrogen and the chemical composition of DNA. The belief, it would seem, was that these concepts were universals—things anyone in the Universe would recognize. *Do Aliens Speak Physics?* asks whether that assumption is warranted.

Suppose we find an alien civilization. Could we talk physics with them at all? That's not really a physics question or a question about aliens. Instead, it's a question for almost every other corner of the sciences and humanities. It's a question about evolutionary biology—about whether very different types of intelligence could arise in the cosmos that would be impossible for humans to communicate with. It's a question about history—about whether humanity could have had technology without science, or science without math, or math with a radically different focus. And it's a question about philosophy—about how sure we are that the laws of the Universe are not just the laws of our own heads.

Whiteson is a physicist, not an evolutionary biologist, historian, or philosopher. Aware of this, he clearly tried to do his homework. The book has an extensive bibliography, and in between illustrations of extraterrestrials that would look at home in Calvin and Hobbes are cartoons of a comprehensive cross section of intellectuals, from Daniel Dennett to Steven Weinberg. This is a book with multiple quotes from Noam Chomsky. (He appears in a wizard robe, profusely thanked for his amazing willingness to answer the authors' emails.)

"Do aliens speak physics?" is thus an extremely ambitious question. And unsurprisingly, the book does not really answer it.

In its attempt, the book considers an extended Drake equation. Building off of Frank Drake's attempt to estimate the number of alien civilizations whose messages we should be able to receive, Whiteson and Warner add terms to calculate how many of those civilizations would be able to talk physics with us, with factors for how likely they are to do something we recognize as science, to be able to communicate their science with us, to be interested in questions we consider scientifically relevant, and to have answers that we can meaningfully learn from. For each factor, the authors spend a couple of chapters trying to break down what sorts of things might contribute to it.

In a scholarly book, a reader would expect to find a probability for each factor, likely with very large error bars. No such estimate appears in this book. That is because *Do Aliens Speak Physics?*, despite pulling in more scholarship on this question than any source I have seen, is still ultimately a pop science book—one that mentions a wide variety of fascinating ideas across disciplines but that, in the name of accessibility, cannot actually delve deeply into any of them.

The best use of *Do Aliens Speak Physics?*, in my mind, is to start a conversation. It is a book you can give to your grandfather to chat about over a long lunch or to your niece to spur a conversation on a summer evening spent on the back porch staring at the stars. Unlike other pop science books, it offers a conversation where neither the reader nor the author is an expert. Instead of a lecture, it is an honest exploration of some of the deepest questions behind science.

I would still love a book that actually answers the question "do aliens speak physics?"—even with the biggest error bars in the cosmos. But lacking that, it is fun to read a book that just asks the question. □

REFERENCES AND NOTES

1. K. Weinersmith, Z. Weinersmith, *A City on Mars: Can We Settle Space, Should We Settle Space, and Have We Really Thought This Through?* (Penguin, 2023).
2. A. Warner, *Spring Rain: A Graphic Memoir of Love, Madness, and Revolutions* (St. Martin's Griffin, 2020).
3. Wikipedia, "Pioneer plaque"; https://en.wikipedia.org/wiki/Pioneer_plaque.
4. Wikipedia, "Arecibo message"; https://en.wikipedia.org/wiki/Arecibo_message.

10.1126/science.aeb1978

The reviewer is a freelance journalist based in Copenhagen, Denmark.
Email: mattvonhippelsciencewriting@proton.me

ILLUSTRATION: ANDY WARNER



In 2023, a strong cold front triggered a super dust event in Haiyuan County, Ningxia Province, China.

PHOTO: NINGXIA NEWS NETWORK

Asian dust threatens air pollution control efforts

In April, a dust storm originating in Mongolia swept more than 3000 km southward over the course of a week, covering eastern China (*1*). Asian dust storms have increased in both intensity and frequency (*2, 3*), and their pathways have shifted to the east and south (*2, 4*). Dust increasingly reaches densely populated regions in eastern China, threatening air quality in major cities (*5*). Over the past decade, China has made substantial progress in reducing urban air pollution (*6*), but large natural particle sources, such as mineral dust, continue to offset air quality improvements (*5*). China should streamline dust and air quality efforts and work with other countries in the region to prevent and manage dust storms.

The China Meteorological Administration monitors dust storms, whereas environmental regulators concentrate on anthropogenic pollution. This regulatory gap leads to an underappreciation of the effects of natural dust, which can decrease air quality and harm human health (*7*). During long-range transport, dust particles undergo complex gas-phase and multiphase chemical reactions, accumulating harmful inorganic and organic pollutants on their surfaces (*8*). As dust storms travel, the particle size of the dust decreases, enhancing toxicity (*9*) and exacerbating health risks in downstream regions, such as Beijing, Shanghai, and Hong Kong.

To maintain its air quality gains, China must prioritize the prevention and management of Asian dust storms. Regional cooperation with upwind source countries, such as Mongolia, is critical. Joint efforts should focus on mitigating the development of dust storms by limiting unsustainable mining and agriculture in vulnerable desert regions and promoting ecological restoration and reforestation (*10*). In addition, China should expand monitoring frameworks, which currently focus on particulate matter smaller than 2.5 μm in diameter (PM2.5) and ozone (O_3) (11), to include routine natural dust storm surveillance and emergency response mechanisms. Policies must account for transboundary dust storms that travel thousands of kilometers and affect air quality in cities far from their source. Public awareness campaigns should emphasize the health risks posed by dust storms and promote protective measures during dust events.

Globally, mineral dust affects the respiratory health of an estimated 330 million people, influences surface O_3 levels, and disrupts ecosystem functions (*5, 12*). Effectively addressing the challenges of Asian dust storms will reinforce China's progress in air pollution control and contribute to global public health, climate adaptation, and sustainable development.

Guochen Wang, Shoujuan Shu, Weijun Li

State Key Laboratory of Ocean Sensing and Department of Atmospheric Sciences, School of Earth Sciences, Zhejiang University, Hangzhou, China. Email: liweijun@zju.edu.cn

REFERENCES AND NOTES

1. World Meteorological Organization, "Severe SDS Sweep Across Much of China" (RSMC-ASDF BEIJING, 14 April 2025); http://www.asdf-bj.net/publish/cms/view/f74171ef822042e7aa38298178ed6666.html.
2. Y. Ma *et al.*, *Global Planet. Change* **242**, 104578 (2024).
3. Z. Yin, Y. Wan, Y. Zhang, H. Wang, *Natl. Sci. Rev.* **9**, nwab165 (2022).
4. Y.-X. Li *et al.*, *Natl. Sci. Rev.* **12**, nwaf274 (2025).
5. Z. Wang, X. Huang, N. Wang, J. Xu, A. Ding, *J. Geophys. Res. Atmos.* **125**, e2020JD033601 (2020).
6. G. Geng *et al.*, *Nat. Geosci.* **17**, 987 (2024).
7. C. Zhang *et al.*, *Nat. Commun.* **14**, 6867 (2023).
8. W. Li *et al.*, *Natl. Sci. Rev.* **12**, nwaf221 (2025).
9. X. Zhang *et al.*, *Atmosphere* **7**, 158 (2016).
10. C. Zucca, R. Fleiner, E. Bonaiuti, U. Kang, *Catena* **219**, 106575 (2022).
11. Q. Xiao *et al.*, *Environ. Sci. Technol.* **56**, 6922 (2022).
12. J. M. Prospero *et al.*, *Global Biogeochem. Cycles* **34**, e2020GB006536 (2020).

10.1126/science.aeb2629

Restore strict protection for wolves in Europe

After centuries of persecution that drove wolf (*Canis lupus*) populations to the brink of extinction in most European countries (*1*), conservation measures and changes in land use have allowed the species to start slowly recovering (*2*). However, on 5 June, the Council of Europe approved the decision of the European Commission (EC) to downgrade the legal status of wolves in the Habitats Directive from Strictly Protected to Protected (*3*). Wolves classified as Strictly Protected could only be removed by derogation and on an individual basis, whereas wolves with Protected status could be removed as a result of culling plans, which puts population recovery in jeopardy (*4*). Given the gaps in knowledge about wolf populations and continued threats to the species in Europe, wolves should remain Strictly Protected.

With the monitoring currently in place, reliably estimating the number of wolves in Europe is not possible. The available demographic data are heterogeneous, approximate, largely qualitative, and unverifiable (*5*). Wolves have not yet reached a favorable conservation status in most European countries (*4*). Up to half of European wolf populations do not meet the effective population size criterion required for long-term genetic and demographic sustainability (*6*).

Wolf populations across Europe are primarily threatened by overhunting (including illegal killing), conflicts driven by livestock depredation, and political pressures from interest groups such as hunters and farmers to reduce population numbers through lethal removal (*4*). New threats are emerging as eastern and southern European countries erect border fences. This growing network of physical barriers, covering thousands of kilometers, has fragmented habitats and disrupted the genetic connectivity of wildlife species, including wolves (*7*).

The proposal to downgrade the wolf's status was based on a non–peer-reviewed technical report commissioned by the EC (*8*) and concerns about wolves' socioeconomic impact (*3*). However, the economic impact of wolf depredation is marginal and can be substantially mitigated by effective prevention programs (*5*), whereas lethal removal has been ineffective (*9*).

LIFE IN SCIENCE



The Coronation Chair of Denmark was crafted from narwhal tusks in the 17th century.

PHOTO: LARS ØIVIND KNUTSEN

A legendary request for narwhal DNA

The Coronation Chair of Denmark, displayed at Rosenborg Castle, a museum in central Copenhagen, is a national treasure that was used in the coronation of Danish kings from 1671 to 1840 (*1*). The throne is pearly white with gold accents, a regal piece befitting its noble history. In 2007, as a PhD student, I called the curator of the museum to request access to the chair. I wanted to drill holes in it.

King Frederik III commissioned the throne in about 1660. According to legend, the chair was made of unicorn horns—rare mythical objects that had healing powers. In fact, the king had sent an expedition to Greenland to collect the spiraled tusks of male narwhals.

I hoped to collect samples from the 350-year-old tusks to investigate why narwhals have lower genetic diversity compared with other marine mammals of the polar north, but I feared that my request would be met with horror. To my astonishment, the curator found the project fascinating and said that my timing was perfect: The throne had been disassembled to undergo restoration.

We arranged for me to visit with my equipment—a dental drill, plastic gloves, and sample tubes. The throne's parts were laid out in a former royal horse stable. Centuries after being cut, the narwhal tusks' distinctive structure had continued to twist into a spiral, so the pieces that formed the backrest and front were soaking in a solution to straighten them. I drilled tiny holes in the large tusk segments that form the throne's pillars. Later, the chair was reassembled, and the holes are now invisible to viewers.

Unlike narwhals in the 17th century, today's narwhal populations face substantial changes to their arctic habitat driven by anthropogenic climate change. The data that I collected could provide information that improves management and conservation policies. I hope that my unusual request will ultimately help safeguard this legendary species.

Eva Garde

Greenland Institute of Natural Resources, Copenhagen, Denmark. Email: evga@ghsdk.dk

REFERENCES AND NOTES

1. J. Hein, Ivories and Narwhal Tusks at Rosenborg Castle: Catalogue of Carved and Turned Ivories and Narwhal Tusks in the Royal Danish Collection 1600–1875, vol. 2 (Museum Tusculanum Press, 2018).

10.1126/science.ady6674

CALL FOR SUBMISSIONS

Life in Science is an occasional feature highlighting some of the humorous or unusual day-to-day realities that face our readers. Can you top this? Submit your story to www.submit2science.org.

To remain consistent with the European Nature Restoration Law, which places ecosystem health above economic and stakeholder interests (*10*), Europe should restore strict protection for wolves and implement reliable population monitoring and policies that recognize their ecological role. Priorities should include cross-species ecosystem recovery, ecological corridors, and livestock depredation mitigation. The EC should integrate science-based information on wolf behavioral ecology into nature conservation programs, considering ecosystem health and resilience as well as human well-being, and should also combat illegal killing and other threats to biodiversity. A full wolf recovery would restore the ability of this top predator to regulate and maintain the health of prey populations and promote reforestation, landscape heterogeneity, and biodiversity, all of which are achieved by both natural and assisted recolonization of wolves in Europe and elsewhere (*11*, *12*).

Ettore Randi[1], Hugh Jansman[2], Josip Kusak[3,4,5], Michela Pacifici[6], Gianluca Piovesan[7], Urmas Saarma[8], Krzysztof Schmidt[9], Geraldine Werhahn[10]

[1]Department of Chemistry and Bioscience, Aalborg University, Aalborg, Denmark. [2]Wageningen Environmental Research, Wageningen, Netherlands. [3]Department of Veterinary Biology, University of Zagreb, Zagreb, Croatia. [4]College of Sciences, Koç University, Istanbul, Turkey. [5]School of Biological Sciences, University of Utah, Salt Lake City, UT, USA. [6]Sapienza University of Rome, Rome, Italy. [7]Department of Ecological and Biological Sciences, University of Tuscia, Viterbo, Italy. [8]Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia. [9]Mammal Research Institute, Polish Academy of Sciences, Białowieża, Poland. [10]Himalayan Wolves Project, IUCN/SSC Canid Specialist Group, Wildlife Conservation Research Unit, University of Oxford, Oxford, UK. Email: ettorerandi17@gmail.com

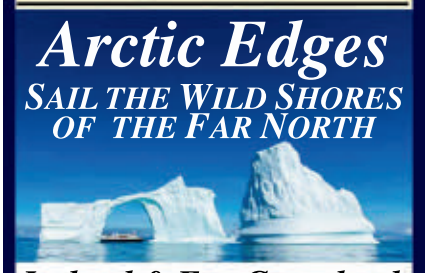
REFERENCES AND NOTES

1. P. Kaczensky *et al.*, "Large carnivore distribution maps and population updates 2017–2022/23. Report to the European Commission under contract no. 09.0201/2023/907799/SER/ENV.D.3" (IUCN/SSC Large Carnivore Initiative for Europe and Istituto di Ecologia Applicata, 2024).
2. L. Boitani *et al.*, "Assessment of the conservation status of the Wolf (*Canis lupus*) in Europe" (Council of Europe Publishing, 2022).
3. Council of the EU, "Habitats directive: Council gives final approval to the new protection status of wolves," press release (5 June 2025); https://www.consilium.europa.eu/en/press/press-releases/2025/06/05/habitats-directive-council-gives-final-approval-to-the-new-protection-status-of-wolves/.
4. M. Kutal, M. Duľa, M. Haring, J. V. López-Bao, *Conserv. Lett.* **18**, e13125 (2025).
5. C. Di Bernardi *et al.*, *PLOS Sustain. Transform.* **4**, e0000158 (2025).
6. J. Mergeay *et al.*, *Evol. Appl.* **17**, e70021 (2024).
7. K. Nowak *et al.*, *Science* **389**, 358 (2025).
8. J. C. Blanco, K. Sundseth, "The situation of the wolf (*Canis lupus*) in the European Union—An in-depth analysis" (N2K Group for DG Environment, European Commission, 2023).
9. R. J. Lennox, A. J. Gallagher, E. G. Ritchie, S. J. Cooke, *Biol. Conserv.* **224**, 277 (2018).
10. E. Chiti, *Eur. L. Open* **3**, 190 (2024).
11. M. RodríguezRecio, C. Wikenros, B. Zimmermann, H. Sand, *Biology* **11**, 317 (2022).
12. W. J. Ripple *et al.*, *Glob. Ecol. Conserv.* **58**, e03428 (2025).

10.1126/science.aeb0660

TRAVEL FOR AAAS
Members and Friends

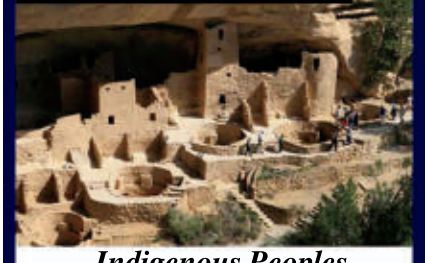
Arctic Edges
SAIL THE WILD SHORES OF THE FAR NORTH



Iceland & East Greenland
August 2 - 9, 2026
Aboard National Geographic Explorer
With 5% Savings!



Arizona Dark Skies
April 18 - 25, 2026



Indigenous Peoples of the Southwest, Arizona in Spring & a Meteor Shower!

5% voyage Savings! From $ 3,876 pp



Wild Alaska Escape
May 11 - 16, 2026

For a detailed brochure, please call (800) 252 - 4910 or (408) 252 - 4910 or Email: Info@BetchartExpeditions.com

BETCHART EXPEDITIONS INC.
17050 Montebello Rd
Cupertino, California 95014
www.BetchartExpeditions.com

McGOVERN INSTITUTE

Call for Nominations: Scolnick Prize in Neuroscience

The McGovern Institute for Brain Research is accepting nominations for the 22nd annual Edward M. Scolnick Prize in Neuroscience. The Prize recognizes an outstanding discovery or significant advance in the field of neuroscience. The prize is $225,000. The recipient presents a public lecture at MIT, hosted by the McGovern Institute and followed by a dinner in Spring 2026.

Nomination Deadline:
December 15, 2025

Nomination procedures:
Candidates for the award must be nominated by individuals affiliated with universities, hospitals, medicals schools, or research institutes, with a background in neuroscience. Self-nomination is not permitted. Each nomination should include:
- A biosketch or CV of the nominee;
- A letter of nomination with a summary and analysis of the major contributions of the nominee to the field of neuroscience.
- Up to two representative reprints will be accepted.

Selection Procedure:
- Members of the selection committee and faculty affiliated with MIT are not eligible.
- Announcement of the award recipient will be made in January 2026.
- Recipient must attend all events to be awarded the prize.

Past Scolnick Prize Recipients:
2025: Leslie Vosshall, HHMI, The Rockefeller University; 2024: Margaret Livingstone, Harvard University; 2023: Yang Dan, HHMI, University of California, Berkeley; 2022: David Ginty, Harvard University; 2020: Joshua Sanes, Harvard University; 2019: Richard Huganir, Johns Hopkins University; 2018: David J. Anderson, HHMI, Caltech; 2017: Catherine Dulac, HHMI, Harvard University; 2016: Cornelia Bargmann, HHMI, The Rockefeller University; 2015: Charles Gilbert, The Rockefeller University; 2014: Huda Zoghbi, HHMI, Baylor University; 2013: Thomas Jessell, HHMI, Columbia University; 2012: Roger Nicoll, UCSF; 2011: Bruce McEwen, The Rockefeller University; 2010: Lily and Yuh-Nung Jan, UCSF; 2009: Jeremy Nathans, Johns Hopkins University; 2008: Michael Davis, Emory University; 2007: David Julius, UCSF; 2006: Michael Greenberg, Children's Hospital/HMS; 2005: Judith Rapoport, NIH; 2004: Masakazu Konishi, CalTech

Send nomination packet to:
gwolf@mit.edu

SPECIAL ISSUE

2025 MANDALAY EARTHQUAKE

ANGELA HESSLER and **SUMIN JIN**

The moment magnitude 7.7 to 7.8 earthquake that struck just outside Mandalay, Myanmar, on 28 March hit the region with strikingly widespread force. Thousands lost their lives, and strong shaking and damage were experienced as far away as Bangkok, Thailand. The research presented in this Special Issue helps to explain the properties and mechanisms behind the rupture and what these mean for earthquake forecasting along the Sagaing and similar faults. Goldberg *et al.* report on the earthquake's anomalously long, fast rupture, which outpaced typical shear-wave velocity and exacerbated ground shaking and damage over much of its ~500-km run. The Mandalay event was also the first time a rupture was captured by a closed-circuit television camera, which was located just meters away. High-resolution video analysis by Latour *et al.* tracked the motion of objects on either side of the fault and found an asymmetric progression of slip over 1.4 s, which provided direct evidence for how pulsed slip added up to produce the full 2-min-long rupture. The sustained, fast rupture may have been enabled by certain properties of the Sagaing fault. Xu *et al.* combined seismic and satellite data to link the rapid transition to supershear rupture to the contrasting rock properties on either side of the fault, and Wei *et al.* found that this transition occurred at a change in fault dip and thickness by integrating multiscale observations to reconstruct the event's displacement field. Together, these studies point to the rupture dynamics and fault properties that drive seismic hazard along this, and other, major strike-slip faults.

10.1126/science.aec3871



PHOTO: CAI YANG XINHUA/EYEVINE/REDUX

Teams from China carry out a search-and-rescue operation in Mandalay, Myanmar, on 1 April after the devastating earthquake.





High-speed fault rupture greatly expanded the damage area in Myanmar.

PHOTO: JIANG CHAO/XINHUA/GETTY IMAGES

2025 MANDALAY EARTHQUAKE

PERSPECTIVE

Mandalay earthquake pushes rupture limits

Fast rupture of Earth's crust breaks a long fault, expanding the damaged area

Kyle E. Bradley[1,2] and **Judith A. Hubbard**[1,2]

On 28 March 2025, a magnitude 7.7 earthquake struck the Sagaing Fault—a fracture in Earth's crust that runs the north-to-south length of Myanmar. The earthquake killed more than 4000 people and damaged a wide region, including the capital (Naypyidaw) and the country's second-largest city (Mandalay). Although a strong earthquake had long been anticipated for this dangerous fault, the rupture that caused the earthquake ran for more than 460 km, which is far longer than expected. On pages 458, 463, 468, and 476 of this issue, Goldberg *et al.* (*1*), Latour *et al.* (*2*), Wei *et al.* (*3*), and Xu *et al.* (*4*), respectively, report the detailed growth of the fault rupture over the course of 2 min. The studies provide a compelling account of the important role of high-speed fault rupture in greatly increasing the area of damage.

Earthquakes are generated by a sudden fault slip: an opposing movement of rock along a fracture in Earth's crust. Strong earthquakes caused by large fault ruptures expose hundreds of millions of people around the globe to intense ground shaking hazard. These earthquakes also provide rare opportunities to study the physical processes that govern natural fault ruptures. Understanding how these ruptures spread along a fault and grow to great size is particularly relevant for the largest earthquakes that cause the greatest harm.

Over the past half century, interest has grown in a type of fault rupture that breaks an apparent speed limit for rupture growth (*5*–*7*). These so-called "supershear" ruptures were once considered to be quite rare. Seismological methods that track rupture growth over time have shown that they are more common than previously

thought (*8*, *9*). However, many questions about the physical processes underlying supershear ruptures, and their implications for seismic hazard, remain unanswered.

The Sagaing Fault is an earthquake "superhighway": a long, straight fault that can host especially fast and dangerous ruptures (*10*). Sections of a fault that have accumulated tectonic strain from constant plate motion but have not released it during recent large earthquakes are called seismic gaps and pose a special hazard. The Ava seismic gap—a 230-km section of the fault between Mandalay and Naypyidaw that had not experienced a major earthquake since 1839—has been particularly worrisome (*11*). Nevertheless, scenarios for supershear rupture during a large earthquake along the Sagaing Fault had been largely speculative.

Goldberg *et al*., Wei *et al*., and Xu *et al*. derived broadly similar models of the evolving fault rupture during the 2025 Mandalay earthquake by combining measurements of seismic waves with satellite measurements of ground movement near the fault. The earthquake nucleated 10 km below Earth's surface and then advanced toward the north and the south along the Sagaing Fault as two separate ruptures. The northbound rupture lasted for only half a minute, growing at a subshear speed (~3 km/s) and causing large fault offsets of up to 6 m. The seismic waves caused by this rupture heavily damaged the Sagaing and Mandalay regions. By contrast, the southbound rupture propagated for almost 2 min, passing Naypyidaw at a supershear speed (~5 km/s) and overrunning the southern end of the Ava seismic gap by 100 km. Sustained supershear rupture was confirmed by the presence of surface wave Mach cones, which are generated when a rupture outpaces its own shear waves. This is analagous to the sonic boom caused by jets flying faster than the speed of sound in air. Wei *et al*. and Xu *et al*. further observed supershear rupture in the spatiotemporal evolution of high-frequency seismic radiation. A seismic recording made near the fault at Naypyidaw also indicated a passing supershear rupture. The total rupture length exceeded 460 km, which is almost twice the length expected of a typical magnitude 7.7 earthquake. These features of the Myanmar earthquake echo the destructive 1906 San Francisco earthquake (magnitude 7.9), which is also thought to have exhibited rupture in two directions, an extremely long rupture length, and a sustained supershear propagation (*12*).

The southbound rupture was fortuitously captured by a security camera located near the fault, providing a high-resolution record of surface slip during an earthquake. Latour *et al*. tracked the surface motions across the fault over time to measure the instantaneous slip rate (the speed at which two fault blocks move past each other) and slip acceleration (the rate of change in slip). At the camera location, the fault slipped for only 1.4 s at a maximum velocity of 3.5 m/s. These measurements show that only a very small area of the large fault was slipping at a given moment. This supports the idea that pulse-like rupture may be common for large crustal earthquakes (*13*).

The studies of Goldberg *et al*., Latour *et al*., Wei *et al*., and Xu *et al*. collectively showed that a sustained supershear rupture can break a much longer section of a fault than predicted by magnitude-length scaling laws. These are observed relationships between the magnitude of an earthquake and the length of a fault rupture and are widely used in seismic hazard modeling. Goldberg *et al*. argued that this excess rupture length doubled the human population exposed to very strong ground shaking compared with a more typical magnitude 7.7 earthquake. The unusual rupture length also implied a shorter-than-expected recurrence time of a successive earthquake on the same fault, suggesting that these events may be more common than previously appreciated. Thus, understanding how and why a sustained supershear rupture occurs is important for preparing for future events.

Numerous physical mechanisms have been proposed to explain the initiation or inhibition of supershear ruptures. These include the material properties of the faulted rocks, complex stresses arising from the fault zone structure, and dynamic pressurization of fluids that fill the pores of rocks inside the fault zone (*14*). Field observations have largely failed to discriminate between these mechanisms and other effects during natural earthquakes. Precisely determining the location and timing of supershear initiation can help reduce this uncertainty but has proved to be difficult.

The inferred initiation location of the Sagaing Fault supershear rupture varied notably across the four studies. Xu *et al*. inferred an almost immediate transition to sustained supershear rupture, implicating the stress conditions near the location where the earthquake began (called the hypocenter). Goldberg *et al*. also suggested an early switch to supershear rupture, followed by a brief subshear interval, and then a second shift to sustained supershear. This series of transitions requires two separate initiation events, each with its own causes. Wei *et al*. determined that the fault rupture initiated at a point located almost 50 km farther south of the hypocenter that was used in Xu *et al*. and Goldberg *et al*. On the basis of this updated hypocenter, Wei *et al*. proposed a single transition to a supershear rupture later in time, suggesting a more important role for the conditions along a larger area of the fault. Latour *et al*. interpret subshear rupture at the camera location, supporting the late transition proposed by Wei *et al*. These disagreements are caused by a lack of near-fault observations, which are critical for pinpointing the hypocenter and tracking rupture behavior (*15*).

In Myanmar, a decade of political liberalization (2011 to 2021) allowed for rapid growth in local efforts to monitor seismic hazards and engineer buildings to withstand earthquakes. Observational networks were also greatly expanded through international collaborations, providing valuable data for earthquake research and providing a training ground for young scientists. After the return of military rule and the onset of civil war in 2021, most of this capacity has disappeared. Field observations, global positioning system (GPS) data, and near-field seismic recordings are thus largely absent for the 2025 Myanmar earthquake. History shows that the period of recovery and assessment after a large earthquake can spur coordinated efforts to reduce the risk of future disasters. This is when new investments in science, engineering, and policy can begin to improve resiliency over the long term. In Myanmar, this period is passing with little governmental action. Although the failed monitoring networks were a missed opportunity for science in 2025, a failed response could represent a far worse missed opportunity for Myanmar for decades to come. □

REFERENCES AND NOTES

1. D. E. Goldberg *et al*., *Science* **390**, 458 (2025).
2. S. Latour *et al*., *Science* **390**, 463 (2025).
3. S. Wei *et al*., *Science* **390**, 468 (2025).
4. L. Xu *et al*., *Science* **390**, 476 (2025).
5. R. Burridge, *Geophys. J. Int.* **35**, 439 (1973).
6. A. J. Rosakis, O. Samudrala, D. Coker, *Science* **284**, 1337 (1999).
7. R. J. Archuleta, *J. Geophys. Res.* **89**, 4559 (1984).
8. H. Bao *et al*., *Nat. Geosci.* **15**, 942 (2022).
9. J. Atterholt, J. J. McGuire, A. J. Barbour, C. Stewart, M. P. Moschetti, *Science* **389**, 1361 (2025).
10. D. P. Robinson, S. Das, M. P. Searle, *Tectonophysics* **493**, 236 (2010).
11. Y. Wang, K. Sieh, S. T. Tun, K. Y. Lai, T. J. Myint, *J. Geophys. Res. Solid Earth* **119**, 3767 (2014).
12. S. G. Song, G. C. Beroza, P. Segall, *BSSA* **98**, 823 (2008).
13. V. Lambert, N. Lapusta, S. Perry, *Nature* **591**, 252 (2021).
14. P. Pampillón, D. Santillán, J. C. Mosquera, L. Cueto-Felgueroso, *Sci. Rep.* **13**, 398 (2023).
15. A. J. Rosakis, M. Abdelmeguid, A. Elbanna, *Nat. Geosci.* **18**, 534 (2025).

10.1126/science.aec0791

[1]Earthquake Insights, Ithaca, NY, USA. [2]Cornell University Department of Engineering, Ithaca, NY, USA. Email: geokyle@gmail.com

RESEARCH ARTICLE

Ultralong, supershear rupture of the 2025 M_w 7.7 Mandalay earthquake reveals unaccounted risk

Dara E. Goldberg[1]*, William L. Yeck[1], Catherine E. Hanagan[2], James Atterholt[1], Haiyang Kehoe[1], Nadine G. Reitman[1], William D. Barnhart[3], David R. Shelly[1], Alexandra E. Hatem[1], David Wald[1], Paul S. Earle[1]

The 28 March 2025 moment magnitude (M_w) 7.7 earthquake in Mandalay, Burma (Myanmar), ruptured 475 kilometers of the Sagaing Fault, which was more than twice the length predicted by magnitude scaling relationships. Kinematic slip models and observation of a Rayleigh Mach wave that passed through parts of Thailand confirmed that rupture occurred at supershear velocities of greater than 5 kilometers per second. The anomalous length exposed a vast population to violent near-fault shaking. The Mandalay earthquake is a modern analog for the M_w 7.9 1906 San Francisco earthquake, another atypically long and fast rupture. Probabilistic seismic hazard analyses use scaling relations that do not account for such long ruptures at moderate magnitudes. This limitation, in conjunction with a likely increased population and infrastructure exposure for atypically long ruptures, contributes to a potential mischaracterization of seismic risk.

On 28 March 2025 at approximately 12:51 local time (6:21 UTC), a moment magnitude (M_w) 7.7 earthquake struck with an epicenter near Mandalay, Burma (Myanmar), the country's second largest city (Fig. 1). According to the US Geological Survey (USGS) community science reporting tool "Did You Feel It?" (DYFI) (*1*), the earthquake was felt in many neighboring countries and reached modified Mercalli intensity (MMI) X (extreme shaking) in near-fault regions. The USGS Prompt Assessment of Global Earthquakes for Response (PAGER) loss product (*2*) indicates that economic loss (estimated ~$53 billion) is likely to exceed the annual gross domestic product of Burma (Myanmar) and that fatalities likely exceeded 10,000, with 18.6 million people experiencing MMI VII (very strong) and stronger shaking (*3*). The earthquake was quickly attributed to the nearby Sagaing Fault, a documented fault capable of hosting large, damaging earthquakes [for example, (*4*, *5*)]. Established magnitude-rupture length scaling relations (*6*–*8*) indicate that the mean fault length of a M_w 7.7 strike-slip earthquake is less than 200 km. Empirical relations are central to earthquake hazard models (*9*) and are used to help estimate shaking intensities. Early DYFI reports of stronger-than-expected shaking and damage south toward Yangon and as far as Bangkok (more than 1000 km from the epicenter) suggested that the fault rupture extended farther than expected according to conventional magnitude-rupture length scaling relations. In the days after the 28 March 2025 earthquake, satellite imagery confirmed that the earthquake rupture extended ~475 km (Fig. 1). This anomalously long rupture inspires a reassessment of the physics of large strike-slip earthquakes and a reevaluation of the magnitude-rupture length scaling relationships used in seismic hazard models.

[1]US Geological Survey, Geologic Hazards Science Center, Golden, CO, USA. [2]US Geological Survey, Earthquake Science Center, Moffett Field, CA, USA. [3]US Geological Survey, Earthquake Hazards Program, Golden, CO, USA. *Corresponding author. Email: degoldberg@usgs.gov

The Sagaing Fault

Burma (Myanmar) lies in a tectonically complex and seismically active region, straddling the India and Eurasia plates (Fig. 1). As the India plate moves northeastward, strain is partitioned with a right-lateral strike-slip component of faulting that is accommodated by the north-south-striking Sagaing Fault (Fig. 1). The Sagaing Fault is roughly 1400 km long and extends northward from the northern Andaman Sea to the Eastern Himalayan Syntaxis (*5*, *10*). The southern Sagaing Fault, the causative fault of the 28 March earthquake, accommodates about 24 mm/year of dextral motion (*4*). The seismogenic depth of the Sagaing Fault is inferred to be variable, generally between 10 and 20 km in depth (*4*, *11*). The Sagaing Fault has hosted many destructive earthquakes, including six magnitude (M) 7 or greater events since 1900. Additionally, a large-magnitude (~M 7.9) earthquake occurred in 1839 in approximately the same location as the 2025 event (Fig. 1) (*5*).

The central to southern Sagaing Fault is commonly subdivided into five along-strike segments (from north to south, the Sagaing, Meiktila, Nay Pyi Taw, Pyu, and Bago segments) (Fig. 1) (*5*). In 2021, interseismic geodetic velocity fields indicated that the northern two segments had accumulated enough slip to host M_w 7.3 and 7.4 earthquakes, respectively, whereas the southernmost segment (Bago) had accumulated the equivalent slip of a M_w 7.3 earthquake (*4*). No comparable interseismic observations were available at the time in the south-central Nay Pyi Taw and Pyu segments. The 2025 Mandalay earthquake spanned from the Sagaing to the Pyu segment and appears to have ruptured through, and beyond, a previously identified ~220-km-long seismic gap between Mandalay and Naw Pyi Daw that had not ruptured in more than a century (*5*).

Evidence for supershear rupture

Pixel tracking of Sentinel-2 optical imagery [supplementary materials (SM), materials and methods] illuminated ~475 km of continuous rupture along the Sagaing Fault from ~22.5°N to ~18.3°N (Fig. 2C), and USGS W-phase moment tensor analysis indicates an ~80-s source duration (37.5 s half duration) (*3*). This duration requires that the roughly 400 km of rupture south of the hypocenter occurred at speeds of ~5 km/s on average, exceeding the local shear-wave velocity. Robinson *et al.* (*12*) identified the Sagaing Fault as an earthquake "superhighway," one of about a dozen faults with long zones largely free from bends and stepovers that might slow or arrest rupture and could therefore allow for rupture speeds exceeding the upper-crustal shear wave speed (~3.5 km/s)—that is, supershear rupture [for example, (*13*, *14*)]. We investigated the details of the rupture process with kinematic rupture modeling of available teleseismic, regional accelerometer observations and the Sentinel-2 pixel tracking results (Fig. 2C).

The finite-fault modeling procedure (*15*) allows for variable rupture velocity and rise time (local rupture duration) throughout the fault, up to 6.5 km/s (SM, materials and methods). From our finite-fault model (Fig. 2), we calculated the approximate rupture velocity v_{rupt} as the distance to the hypocenter (d_{source}) divided by the sum of allowable onset time at each subfault (t_{onset}) and half the rise time (t_{rise}) (Fig. 3D)

$$v_{rupt} = \frac{d_{source}}{t_{onset} + {}^{t_{rise}}/_{2}} \quad (1)$$

The denominator of Eq. 1 is the time to peak moment release on the subfault, which we believe to be a more robust measurement than onset time alone because our modeling does not require slip to occur at the earliest allowable onset time. Therefore, our calculated rupture velocity is a minimum estimate of the true velocity of the rupture front. Furthermore, because we calculated average rupture velocity from the hypocenter, or location of rupture onset, instantaneous velocities would necessarily reach even greater values. Rupture propagated to the north

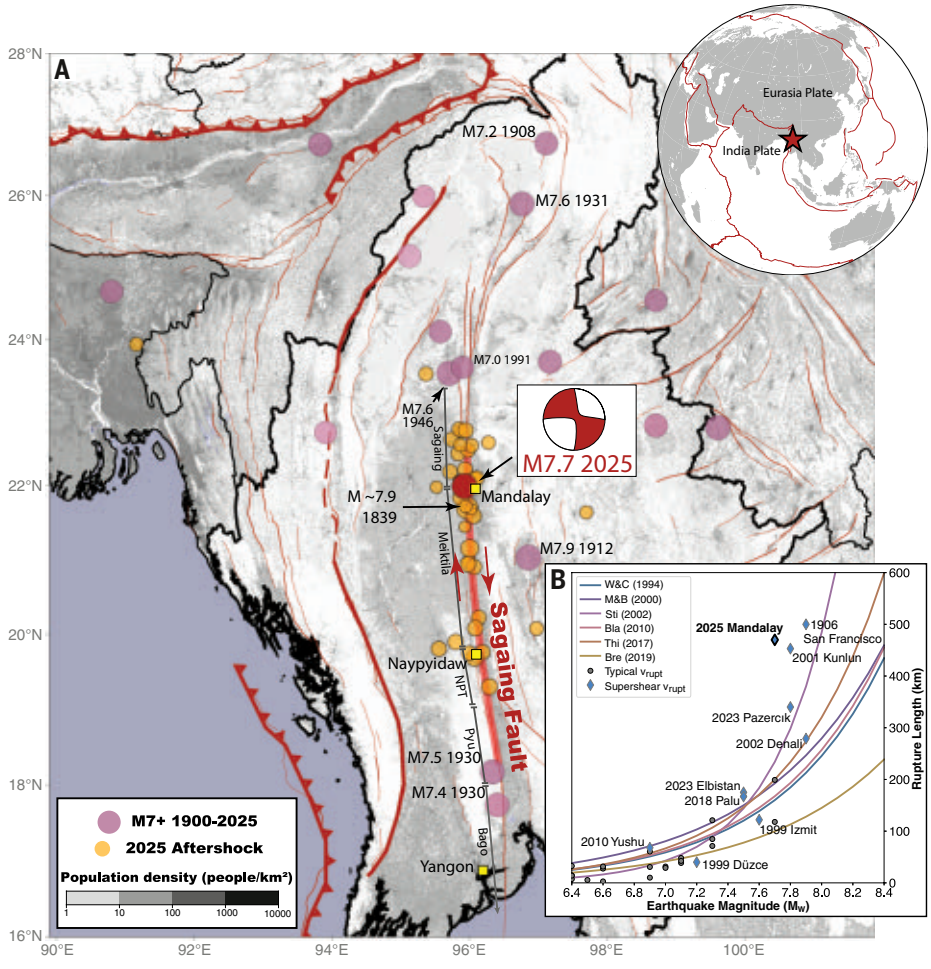


Fig. 1. Tectonic overview and earthquake length scaling. (**A**) Historic *M* 7 or greater seismicity dating back to 1900 is shown as pink circles. An 1839 earthquake (estimated *M* ~7.9) is shown at nearly the same location of the 2025 M_w 7.7 earthquake (red circle and red W-phase centroid moment tensor). The ~475-km-long rupture length of the 2025 mainshock is indicated in red along the Sagaing Fault. Aftershocks of the 2025 mainshock are indicated with orange circles. The populous cities of Mandalay, Naypyidaw, and Yangon are indicated with yellow squares. Major tectonic boundaries are indicated in dark red, with smaller mapped faults as thin red lines (*29*). Labeled gray lines indicate the approximate bounds of the five segments of the southern Sagaing Fault. The basemap shows population density (*30*). (Inset) Major tectonic boundaries as red lines, with the location of the 2025 M_w 7.7 earthquake indicated with the red star. (**B**) Comparison of fault rupture lengths from ruptures known to have any period of supershear rupture (blue diamonds) and sub- to shear-velocity ruptures (gray circles) (*31*–*33*) with common earthquake scaling relationships for continental strike slip faults. W&C (1994), M&B (2000), Sti (2002), Bla (2010), Thi (2017), and Bre (2019) refer to (*6*), (*22*), (*23*), (*7*), (*24*), and (*8*), respectively.

largely at subshear-wave velocities (Fig. 3D); however, the rupture front in the southern propagation direction quickly reached supershear speeds of at least 5 km/s before briefly slowing to shear speeds (~3.5 km/s) about 100 km south of the epicenter. From there, supershear velocities resumed until ~325 km south of the epicenter before decreasing to roughly shear wave speeds beyond 330 km south of the epicenter (Fig. 3D). The modeled source-time function determined the rupture duration to be ~85 s, with later moment attributed to fitting waveform noise (Fig. 2B). Because we calculated a lower bound on the maximum rupture speed, some areas indicate rupture velocity in excess of the Eshelby speed (~5.0 km/s for a shear wave velocity of 3.5 km/s). Our model indicates that more than 53% of the seismic moment of this earthquake occurred before the arrival of the shear wave from the hypocenter; south of the hypocenter, this fraction was more than 68%.

Numerical modeling suggests that supershear rupture can result from vertically polarized shear wave (SV) to P-diffracted wave conversions at the free surface (*16*). This mechanism leads to rupture velocities that are greatest near the surface, despite crustal seismic velocities being lesser near the surface than at depth. As shown in Fig. 2A, we observed this effect prominently in our rupture model, with the shallow rupture front persistently preceding deeper rupture during supershear propagation south of the hypocenter. This effect may also lead to surface-slip displacements that are larger than slip at depth, in contrast to a coseismic shallow slip deficit that is commonly observed in earthquake ruptures (*17*). This pattern can be observed in Fig. 3A, in which measured surface rupture (*18*) and modeled shallow slip exceeds slip at medium depths in places where rupture speeds are high, particularly between 110 and 200 km south of the hypocenter and again between 250 and 350 km south of the hypocenter. Still, slip at medium depth exceeds shallower slip at the north end of rupture, where subshear to shear-velocity rupture occurred. Generally, measured surface rupture and modeled shallow (<7 km) slip follow similar trends, which provides confidence in the model. Perfect correlation between the two is not expected because of the model's depth-averaging of slip over each subfault, model smoothing, and the lack of guarantee that the surface slip displacements are fully coseismic.

Our modeling of supershear rupture is confirmed by direct observation of a Rayleigh Mach wave, which is analogous to the sonic boom generated by objects that exceed the speed of sound in the atmosphere (SM, materials and methods). Theory predicts that surface waves from all sections of a supershear segment will constructively interfere in Mach cones, the angle of which is dependent on the supershear rupture velocity and the Rayleigh wave phase velocity (*19*). This interference results in surface waves that lack the complexity typically observed for large earthquakes, which would be considered as line sources. At these azimuths, surface-wave waveforms of the mainshock instead match scaled waveforms of smaller aftershocks that can be considered as point sources. By comparing the distribution of Rayleigh waves generated by the mainshock with those of multiple large aftershocks (*20*, SM, materials and methods), we confirmed the presence of a Rayleigh Mach wave (Fig. 4).

Areas of faults that undergo supershear rupture often have a dearth of aftershocks, with aftershocks preferentially occurring on adjacent structures or in subshear velocity portions of the fault (*21*). For the first 16 days after the rupture, we observed a paucity of aftershocks in supershear zones (Fig. 3C). We included earthquakes located within 20 km of the rupture trace, with the understanding that more precise relocations might confirm these to be on the same fault as that of the main rupture. We found denser aftershock cluster along the Sagaing Fault in the epicentral region and northward, the region that experienced subshear velocity rupture, as compared with south of the epicenter, where rupture reached supershear velocities. A small cluster of four aftershocks between ~90 and 130 km south of the epicenter

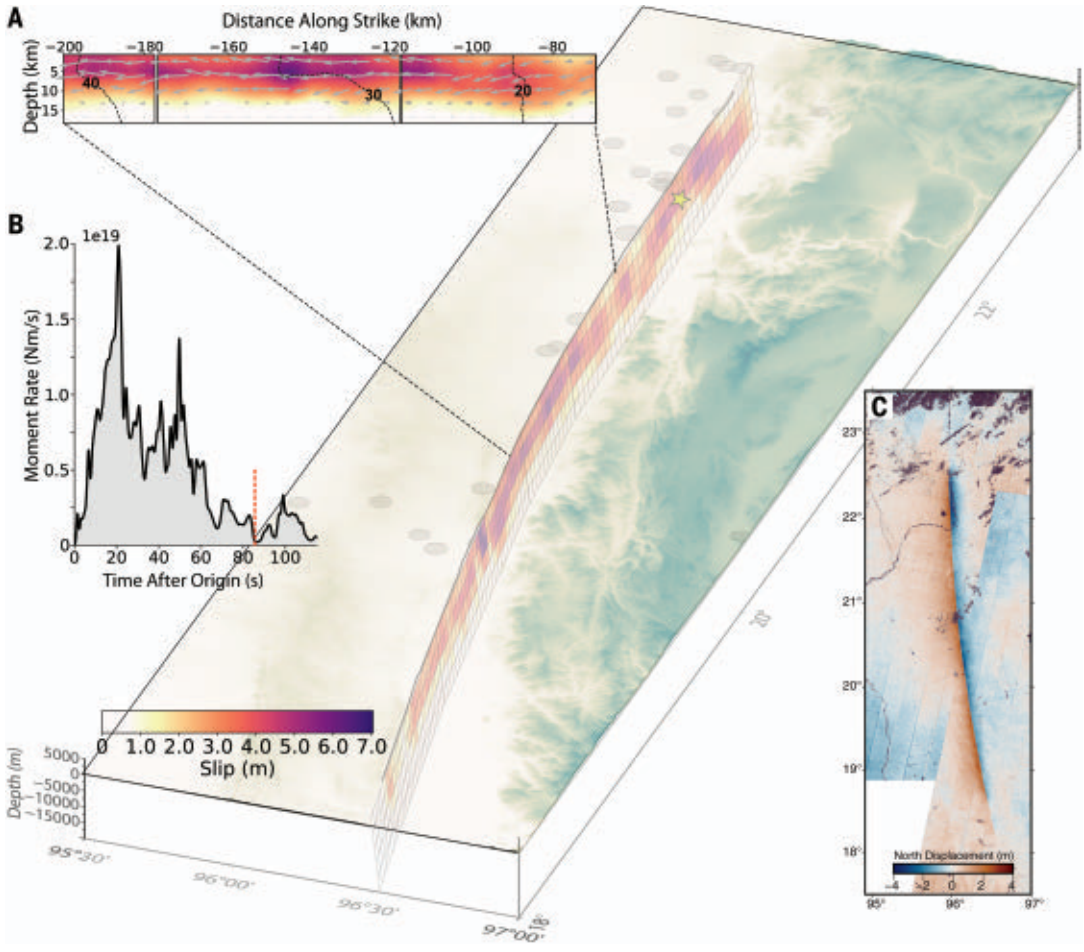


Fig. 2. Finite-fault model. Perspective view of the four-segment discretized fault plane is shown to align with a simplified fault trace (black line). The yellow star indicates the hypocenter location. Aftershock locations are indicated with gray circles. The color bar denotes amplitude of slip. Basemap topography is from (*34*). (**A**) Selected portion of the rupture shown in fault-perpendicular view, with slip smoothed across subfaults. Color indicates slip amplitude as on the central image. Gray arrows indicate the rake direction, scaled by amplitude of slip. Black dashed lines indicate the rupture onset time, labeled in increments of 10 s from 20 to 40 s after origin time. (**B**) The source time function, with slip after 85 s (dashed red line) attributed to fitting waveform noise. (**C**) Pixel-tracking displacements in the north-south direction (primary direction of motion) from Sentinel-2 orbit 46. Positive displacements (red) indicate northward motion; negative displacements (blue) indicate southward motion. Model fits to teleseismic body waves, teleseismic surface waves, local strong-motion accelerometer observations, and Sentinel-2 optical pixel tracking displacements are available in figs. S2, S3, S4, and S5 to S8, respectively.

(Fig. 3C) potentially corresponds to a small region of slower rupture velocity seen in the finite-fault results in the same region (Fig. 3D).

Is the Mandalay rupture anomalous?

The Mandalay earthquake stands out among continental strike slip earthquakes with well-documented rupture lengths (Fig. 1B). The earthquake exhibited one of the longest documented surface ruptures. With respect to earthquake scaling relationships (*6*–*8*, *22*–*24*), the Mandalay earthquake is a clear outlier among continental ruptures, comparable only with the ~450-km-long 2001 M_w 7.8 Kunlun and the ~500-km-long 1906 M_w 7.9 San Francisco earthquakes (Fig. 1B). Most other documented supershear ruptures do not strongly deviate from published magnitude-length scaling relationships (Fig. 1B). Moreover, our conservative estimate that ~65% of the southern moment release occurred during supershear rupture indicates that this rupture velocity was sustained in a manner not typical of other documented supershear ruptures. Thus, the extreme length of the Mandalay earthquake is likely related to the persistence of supershear rupture throughout the earthquake. The sustained high rupture velocities allowed the earthquake to propagate over a longer distance as compared with that of other M 7 or greater earthquakes, thus increasing exposure of population centers along the Sagaing Fault to near-source strong shaking. USGS PAGER reports based on a ~200-km-long rupture (the expected rupture length from scaling relationships) indicated that ~7.7 million people were exposed to MMI VII (very strong) and stronger shaking (PAGER version 6) (*3*). This exposure more than doubled to 18.6 million people when the underlying shaking model included the full 475-km-long rupture (PAGER version 22) (*3*). As an example, the city of Nay Pyi Taw (population ~925,000) was expected to experience strong shaking for a ~200-km-long rupture, with moderate damage to vulnerable structures but light damage to resistant structures (PAGER version 6). By the time the full extent of rupture was understood, Nay Pyi Taw was expected to have experienced violent shaking, accompanied by heavy damage to even resistant structures (PAGER version 22).

Seismic hazard considerations for continental strike-slip fault systems

Satellite imagery indicates that the Mandalay earthquake ruptured the Sagaing, Meiktila, and Nay Pyi Taw fault segments and at least partially ruptured the Pyu segment. On the basis of the earthquake magnitude (M_w 7.7) and previous estimates of accumulated slip (*4*), we

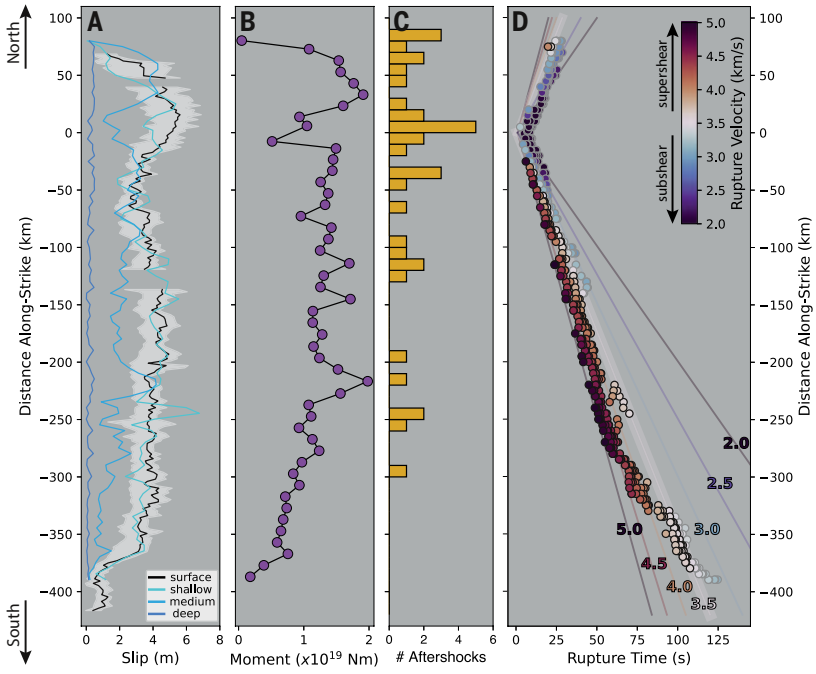


Fig. 3. Fault slip and approximate rupture velocity. (**A**) Slip as a function of depth. Surface slip was measured from satellite observations (*18*), where the light gray area indicates the range of minimum to maximum estimated surface slip, with the average value indicated in black. Modeled shallow (<7 km), medium (7 to 16 km), and deep (>16 km) slip are indicated with shades of blue. (**B**) Moment along-strike, as indicated by the finite fault model solution. Circles indicate measurement in 10-km bins along-strike, connected with black lines for visual clarity. (**C**) Histogram of aftershocks, within 16 days and 20 km of the fault, along-strike. (**D**) Rupture time versus distance along-strike for each subfault (circles) in the finite fault model. Each point is colored by the average rupture velocity from the hypocenter (Eq.1). Colored lines indicate different rupture velocities, with 3.5 km/s indicated with a thicker line to emphasize the approximate shear wave velocity.

can surmise that the Nay Pyi Taw and Pyu segments together stored the interseismic moment of a ~M_w 7.4 earthquake. However, in absence of interseismic strain accumulation estimates in these segments, we cannot discern whether the 2025 earthquake fully or partially ruptured these segments. Because the Bago segment was unruptured in this earthquake, previously reported strain conditions of this segment may remain (*4*).

The longer-than-expected rupture length observed in the 2025 Mandalay earthquake has important implications for probabilistic seismic hazard analysis and impact assessment. Analysis based on established scaling relationships [for example, (*6*)], reported by Wang *et al.* (*5*) determined that the ~220-km-long Meiktila and ~180-km-long Sagaing sections of the Sagaing Fault could host earthquakes of M_w 7.9 and M_w 7.7, respectively, but their combined rupture (totaling ~400 km) could generate a M_w 8.1 to 8.3 earthquake. These authors suggested that return times of such earthquakes would be on the order of 500 to 1000 years. However, the 2025 event ruptured ~475 km [equivalent to more than three fault sections, as defined by Wang *et al.* (*5*)] and did so with a much smaller total seismic moment, equivalent to M_w 7.7. Although a M_w 7.7 may be less damaging than a potential M_w 8 or greater event over the same area, the shorter recurrence interval of smaller events may make these lesser-magnitude events more influential in the magnitude-frequency relationships used in probabilistic seismic hazard analyses.

The Mandalay earthquake is an important analog for other major earthquakes, such as the 1906 M_w 7.9 San Francisco earthquake (*25*, *26*). California's San Andreas Fault system, which hosted the 1906 earthquake, is tectonically similar to the Sagaing Fault; both are

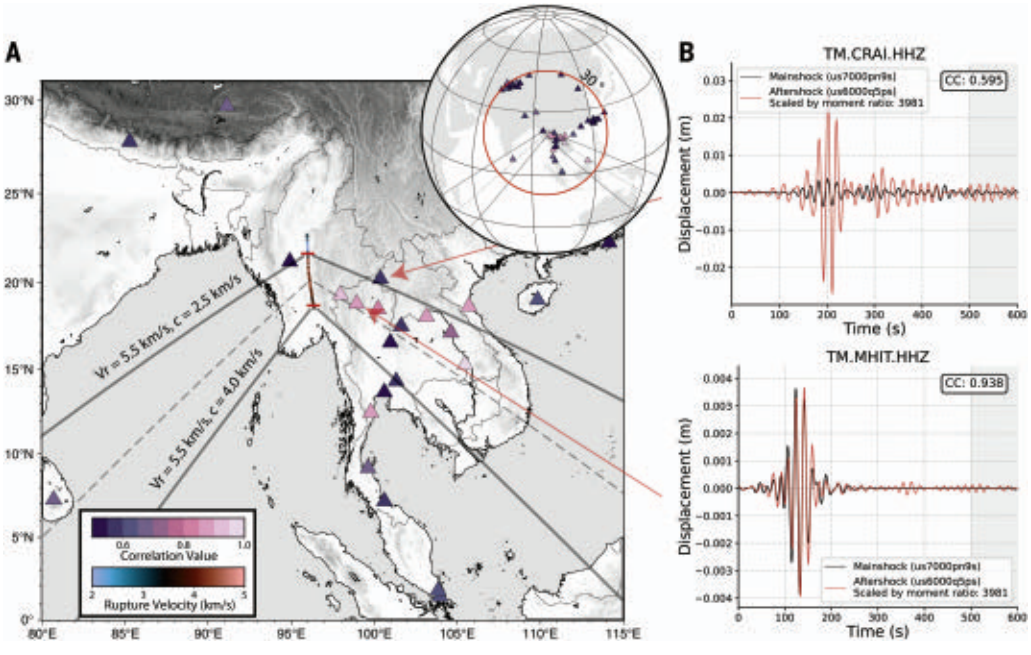


Fig. 4. Mach wave analysis. (**A**) Map view of study region with fault colored by surface rupture velocity. Basemap topography is from (*34*). The Rayleigh Mach wave angles assuming phase velocities (*c*) of 2.5 and 4.0 km/s and a rupture velocity (V_r) of 5.5 km/s are indicated with gray lines. The top and bottom of the supershear rupture are indicated with red lines. Broadband seismic stations are indicated with triangles, colored by the waveform average correlation value (CC) with the selected aftershock. (Inset) A larger view of the seismic stations considered, on the same color scale. (**B**) Example waveforms from stations (top) TM.CRAI and (bottom) TM.MHIT. At station TM.CRAI, outside the Rayleigh Mach wave regions, mainshock waveforms (black) have low correlation with the lesser-magnitude event (red). Station TM.MHIT, in the Mach region, has high correlation. Aftershock waveforms (red) are scaled by 3981, derived from the moment ratio of the mainshock to aftershock. USGS event page identification numbers for the mainshock and aftershock are shown parenthetically in the legends.

right-lateral, geometrically simple continental faults capable of hosting M 8 or greater earthquakes. The 1906 earthquake produced a similar anomalously long surface rupture (~500 km) (Fig. 1B) (*25*, *26*) where earthquake scaling relationships predicted a ~220-km-long rupture. The 1906 earthquake additionally likely ruptured with sustained supershear velocities north of San Francisco (*26*).

The San Andreas Fault also shares societal similarities to the Sagaing Fault. Both pierce major cities with large populations and critical infrastructure. Earthquake rupture forecasts for the San Andreas Fault generally follow common magnitude scaling laws (*27*); these scaling laws are used with a single prescribed seismogenic width, which likely limits the occurrence of extremely long (high aspect ratio) ruptures such as the 2025 Mandalay earthquake. Of the ≤1-million-year recurrence interval right-lateral ruptures of M_w 7.65 to 7.75 considered in the USGS National Seismic Hazard Model 2023, the longest rupture is 382 km long (*28*), which is nearly 100 km shorter than the 2025 Myanmar rupture.

However, as more supershear ruptures are identified with associated atypical source characteristics, it becomes clear that seismic hazard analyses may be incomplete without these outlier events. Although we do not speculate here on the net effect that such outlier events have on overall modeled seismic hazard, the larger exposure area of an outlier event could potentially lead to a broader economic and humanitarian impact because of the quantity of affected infrastructure and population. The 2025 Mandalay earthquake affected multiple populous cities simultaneously, highlighting that the exposure area of lesser-magnitude, sustained supershear ruptures can be comparable with that of an average, greater-magnitude event. Because lesser-magnitude events are more frequent than their greater-magnitude counterparts, quantifying the probability of anomalously long, lesser-magnitude events is critical for constraining overall risk. The range of impacts estimated by the USGS PAGER system, which varied as a function of inferred rupture length, underscores the continued need for resources to rapidly estimate the true and full extent of earthquake shaking to support efficient and effective humanitarian relief.

REFERENCES AND NOTES

1. D. J. Wald, V. Quitoriano, C. B. Worden, M. Hopper, J. W. Dewey, *Ann. Geophys.* **54**, 5 (2012).
2. D. J. Wald, K. S. Jaiswal, K. D. Marano, D. B. Bausch, M. G. Hearne, *U.S. Geol. Surv. Fact Sheet* **2011**, 2010–2036 (2011).
3. M. G. Hearne, libcomcat, version 2.1.7. *US Geological Survey software release* (2025); https://doi.org/10.5066/P91WN1UQ.
4. T. Z. H. Tin *et al.*, *J. Asian Earth Sci.* **228**, 105125 (2022).
5. Y. Wang, K. Sieh, S. T. Tun, K. Y. Lai, T. Myint, *J. Geophys. Res. Solid Earth* **119**, 3767–3822 (2014).
6. D. L. Wells, K. J. Coppersmith, *Bull. Seismol. Soc. Am.* **84**, 974–1002 (1994).
7. L. Blaser, F. Krüger, M. Ohrnberger, F. Scherbaum, *Bull. Seismol. Soc. Am.* **100**, 2914–2926 (2010).
8. C. M. Brengman, W. D. Barnhart, E. H. Mankin, C. N. Miller, *Bull. Seismol. Soc. Am.* **109**, 1701–1715 (2019).
9. M. D. Petersen *et al.*, *Earthq. Spectra* **40**, 5–88 (2024).
10. C. T. Mon *et al.*, *Geophys. Res. Lett.* **47**, e2019GL086236 (2020).
11. W. Fadil *et al.*, *Bull. Seismol. Soc. Am.* **113**, 613–635 (2023).
12. D. P. Robinson, S. Das, M. P. Searle, *Tectonophysics* **493**, 236–243 (2010).
13. A. Socquet, J. Hollingsworth, E. Pathier, M. Bouchon, *Nat. Geosci.* **12**, 192–199 (2019).
14. H. Bao *et al.*, *Nat. Geosci.* **12**, 200–205 (2019).
15. P. Koch *et al.*, Wavelet Inversion for Slip (WISP), version 1.1.0. US Geological Survey software release (2025); https://doi.org/10.5066/P14RVF65.
16. Y. Kaneko, N. Lapusta, *Tectonophysics* **493**, 272–284 (2010).
17. J. M. Nevitt *et al.*, *Sci. Rep.* **10**, 5031 (2020).
18. N. G. Reitman *et al.*, Remote surface rupture observations for the M7.7 2025 Mandalay, Burma (Myanmar) earthquake. US Geological Survey data release (2025); https://doi.org/10.5066/P1RYMWCK.
19. M. Vallée, E. M. Dunham, *Geophys. Res. Lett.* **39**, 2011GL050725 (2012).
20. W. L. Yeck, neic-supershear-analysis, version 1.0.0. US Geological Survey software release (2025); https://doi.org/10.5066/P1BQPVIV.
21. M. Bouchon, H. Karabulut, *Science* **320**, 1323–1325 (2008).
22. P. M. Mai, G. C. Beroza, *Bull. Seismol. Soc. Am.* **90**, 604–615 (2000).
23. M. Stirling, D. Rhoades, K. Berryman, *Bull. Seismol. Soc. Am.* **92**, 812–830 (2002).
24. K. K. S. Thingbaijam, P. M. Mai, K. Goda, *Bull. Seismol. Soc. Am.* **107**, 2225–2246 (2017).
25. W. Thatcher, G. Marshall, M. Lisowski, *J. Geophys. Res.* **102** (B3), 5353–5367 (1997).
26. S. G. Song, G. C. Beroza, P. Segall, *Bull. Seismol. Soc. Am.* **98**, 823–831 (2008).
27. B. E. Shaw, *Bull. Seismol. Soc. Am.* **113**, 924–947 (2023).
28. E. H. Field *et al.*, *Bull. Seismol. Soc. Am.* **114**, 523–571 (2023).
29. E. Zelenin, D. Bachmanov, S. Garipova, V. Trifonov, A. Kozhurin, *Earth Syst. Sci. Data* **14**, 4489–4503 (2022).
30. V. Lebakula *et al.*, *Sci. Data* **12**, 495 (2025).
31. A. Sarmiento *et al.*, Flatfiles and related digital products for the fault displacement hazard initiative database, version 1. UCLA Dataverse (2024); https://doi.org/10.25346/S6/Y4F9LJ.
32. H. Bao *et al.*, *Nat. Geosci.* **15**, 942–949 (2022).
33. M. Abdelmeguid *et al.*, *Commun. Earth Environ.* **4**, 456 (2023).
34. B. Tozer *et al.*, *Earth Space Sci.* **6**, 1847–1864 (2019).

ACKNOWLEDGMENTS

We thank the USGS National Earthquake Information Center (NEIC) 24/7 Duty Seismologists who are critical to the USGS NEIC's global earthquake monitoring. We also acknowledge R. Briggs, W. Mooney, R. Gold, B. Shiro, K. Marra, A. Hesser, and three anonymous reviewers for their valuable comments. **Funding:** The US Geological Survey Earthquake Hazards Program supported this work. **Author contributions:** Conceptualization: D.E.G., W.L.Y., W.D.B., D.R.S., P.S.E. Methodology: D.E.G., W.L.Y., C.E.H., J.A., H.K., N.G.R., W.D.B. Investigation: D.E.G., W.L.Y., C.E.H., J.A., H.K., N.G.R., W.D.B., D.R.S., A.E.H., P.S.E., D.W. Visualization: D.E.G., W.L.Y., W.D.B. Writing – original draft: D.E.G., W.L.Y., C.E.H., J.A., H.K., N.G.R., W.D.B., D.R.S., A.E.H., P.S.E., D.W. Writing – review & editing: D.E.G., W.L.Y., C.E.H., J.A., H.K., N.G.R., W.D.B., D.R.S., A.E.H., P.S.E., D.W. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The USGS event webpage contains up-to-date information of USGS event characterization (https://earthquake.usgs.gov/earthquakes/eventpage/us7000pn9s/executive), including origin information, moment tensor modeling, finite-fault modeling, DYFI, and PAGER. Earlier versions of these products (for example, PAGER) can be accessed using libcomcat (*3*). The facilities of EarthScope Consortium were used for access to waveforms and related metadata for modeling performed outside of NEIC's real-time system (*4*). These services are funded through the Seismological Facility for the Advancement of Geoscience (SAGE) Award of the National Science Foundation under Cooperative Support Agreement EAR-1851048. Sentinel-2 observations are Copernicus Sentinel data processed by the European Space Agency (ESA). The finite fault modeling software is available at https://doi.org/10.5066/P14RVF65 (*15*). The Mach wave detection software is available at https://doi.org/10.5066/P1BQPVIV (*20*), A tagged version of this Mach wave software used can be found at https://code.usgs.gov/ghsc/neic/research/neic-supershear-analysis/-/tags/version_in_publication. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US government. **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ady3581
Materials and Methods; Figs. S1 to S8; References (*35*–*44*)

Submitted 18 April 2025; resubmitted 16 June 2025; accepted 19 August 2025
10.1126/science.ady3581

RESEARCH ARTICLE

Direct estimation of earthquake source properties from a single CCTV camera

Soumaya Latour[1]†, Mathias Lebihain[2], Harsha S. Bhat[3], Cédric Twardzik[4], Quentin Bletery[4], Kenneth W. Hudnut[5], François Passelègue[4]*†

We present a direct measurement of the slip-rate function from a natural coseismic rupture, recorded on 28 March 2025, during the moment magnitude (M_w) 7.7 Mandalay earthquake (Myanmar). This measurement was made using video footage of the surface rupture captured by a closed-circuit television (CCTV) security camera located only meters away from the fault trace. Using direct image analysis, we measured the relative slip at each time step and deduced the slip rate. Our results show a local slip duration of 1.4 seconds and a cumulative slip of ~3 meters, during which surface slip velocity peaked at ~3.5 meters per second with passage of the rupture front. These findings demonstrate the pulse-like nature of the seismic rupture at the location of the recording. Using slip-pulse elastodynamic rupture models, we obtained the complete mechanical properties of this pulse, including the energy release rate.

Estimating the spatial and temporal evolution of slip along fault interfaces is critical for understanding the physics of deformation processes in Earth's crust throughout the seismic cycle (*1*). Because slip typically occurs at depth under extreme conditions, direct in situ observations are, in most cases, impossible. Reconstructions of slip history rely on inverse modeling (*2*). Consequently, our understanding of earthquake physics is fundamentally limited by the resolution and coverage of the data used in these inversions, as well as the inherent complexity of the forward problem (*3*–*6*). As a result, many key aspects of earthquake rupture, particularly the local dynamics of slip and the associated stress evolution at the fault interface, remain poorly constrained.

On 28 March 2025, a devastating moment magnitude (M_w) 7.7 earthquake struck Myanmar along the Sagaing fault (*7*) near Mandalay. The earthquake caused catastrophic damage: more than 5400 fatalities, >11,000 injuries, and hundreds to thousands reported missing. Infrastructure losses included 120,000 homes, 2500 schools, numerous temples, and key transportation networks such as bridges and airports. Several historic sites, particularly in Inwa, suffered extensive damage. This right-lateral strike-slip event ruptured more than 450 km of the fault, with slip reaching the surface on long segments and horizontal displacements up to 6 m [Fig. 1A; (*8*)]. The moment source function derived from the US Geological Survey (USGS) (*9*) indicates that the rupture propagation lasted ~120 s (Fig. 1B), in agreement with the SCARDEC solution (*10*), and likely included supershear phases (*11*).

Beyond its societal impact, the earthquake offered an unprecedented scientific opportunity: For the first time, a closed-circuit television (CCTV) security camera located just meters from the surface rupture captured the real-time deformation of the ground during seismic faulting (*12*). This footage provides a direct observation of coseismic surface displacement, enabling us to extract the time history of slip and slip rate at a specific fault point, measurements that, until now, were only accessible through laboratory experiments or inferred indirectly through modeling. In this study, we analyzed the footage to obtain a direct measurement of the slip-rate function during a natural earthquake (*13*) and, additionally, to invert the evolution of shear stress, energy dissipation (breakdown work), and rupture dynamics. Our results provide ground-truth constraints on seismic rupture physics, offering a critical benchmark for validating numerical models and seismic inversions and helping to close a long-standing observational gap in earthquake science.

Measurement of local slip rate during the earthquake

The camera was positioned on the east side of the fault and was approximately oriented in the southwest direction. To estimate the local slip rate, we tracked distinct visual landmarks located on both sides of the fault (Fig. 2A). The primary landmark was a concrete or metallic pillar ($R0$) located on the west side of the fault; two additional landmarks, $R1$ and $R2$, were small poles located adjacent to a road and behind a fence composed of vertical bars. Throughout the video, these features display predominantly horizontal motion relative to the foreground, consistent with right-lateral strike-slip faulting.

To quantify this motion, we defined two fault-parallel lines intersecting the vertical landmarks (Fig. 2A) and tracked the pixel intensity along these lines over the entire duration of the observable rupture propagation (Fig. 2, B to E). We defined $t = 0$ s as the approximate onset of fault slip [corresponding to the time stamp 12:46:34.633 of the video and not synchronized to universal time coordinated (UTC)]. Ground shaking began at $t = -4.5$ s, which is evident as coherent motion across the frame due to camera motion. A secondary seismic phase was observed at approximately $t = -1.8$ s. The onset of fault slip was identified as the moment when the tracked landmarks began to move relative to nearby, foreground features. This relative motion continued until approximately $t = 1.6$ s (Fig. 2C).

For each landmark, we selected a nearby visually stable reference point on the east side of the fault, as shown in Fig. 2A: $R0'$ (a vertical wall) and $R1'$ and $R2'$ (vertical bars of the fence). We tracked the motion of both the landmark and its corresponding reference feature over time (Fig. 2, C and E). The difference in their pixel positions provided a measure of relative displacement that was corrected for camera motion. We determined the temporal and spatial resolutions using the method described in the supplementary materials (fig. S1), and we applied a spatial scaling using an estimate of the final slip. This approach yielded a temporal resolution of 0.033 s and a spatial resolution with a relative uncertainty of approximately ±20%.

The slip history derived from the three measurements is shown in Fig. 3A. The curve obtained from $R0$ provides the most reliable estimate, as both the onset and arrest of motion were clearly visible in the video. The timing and slope of estimates from $R1$ and $R2$ are consistent with that of $R0$. The slip increased smoothly from 0 to 3 m over ~1.4 s, after which it ceased. We obtained the slip-rate function (Fig. 3B) by numerically differentiating the slip history after resampling it at 90 Hz. We then applied a moving average with a 0.33-s time window, as we interpreted the oscillations in the raw curve to result from uncertainties in manual tracking and camera-motion correction. The slip-rate function is slightly asymmetric: The slip rate rose rapidly from 0 to 3.5 m s^{-1} within 0.6 s, then decreased more gradually to 0 m s^{-1} over ~1 s. These values correspond to an average slip acceleration of 5.8 m s^{-2} and a slip deceleration of 3.5 m s^{-2}.

The total slip duration is best estimated using the rise time of the slip function, as the smoothing applied during signal processing tends to artificially broaden the slip-rate pulse. The most robust estimate of

[1]Université de Toulouse, CNRS, Observatoire Midi-Pyrénées, IRAP, Toulouse, France. [2]Navier, ENPC, Institut Polytechnique de Paris, Université Gustave Eiffel, CNRS, Marne-la-Vallée, France. [3]Laboratoire de Géologie, École Normale Supérieure, CNRS, UMR 8538, PSL Université, Paris, France. [4]Université Côte d'Azur, CNRS, Observatoire de la Côte d'Azur, IRD, Géoazur, Sophia Antipolis, France. [5]Southern California Edison, Rosemead, CA, USA. ***Corresponding author. Email: francois.passelegue@cnrs.fr** †These authors contributed equally to this work.

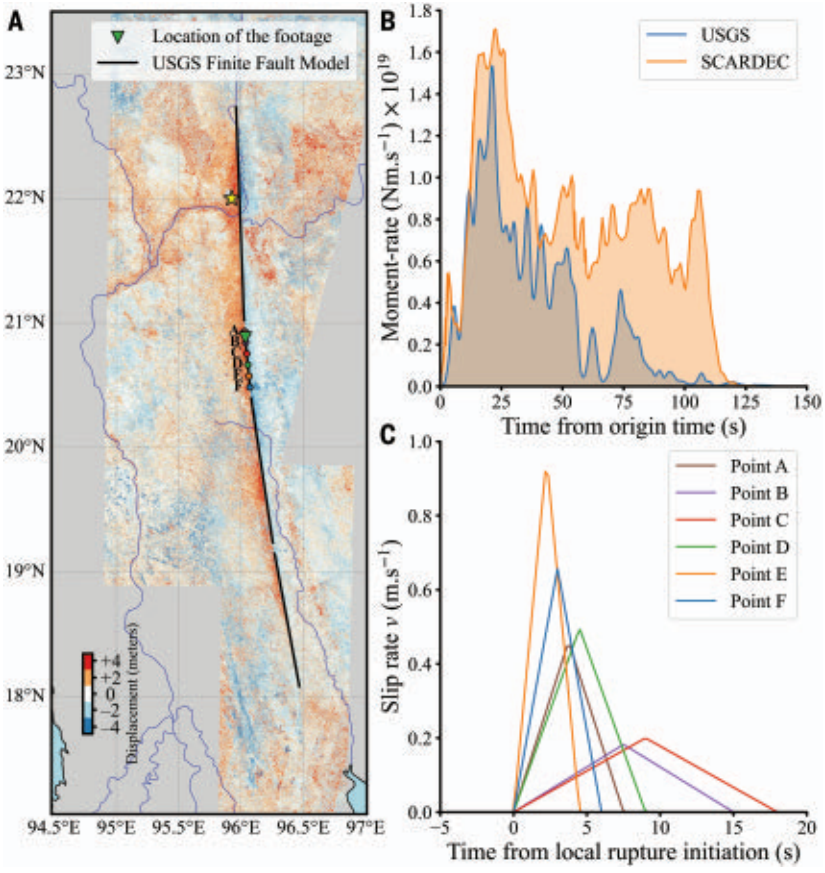


Fig. 1. Displacement field, moment-rate functions, and local slip-rate functions for the 2025 Mandalay earthquake. (**A**) North-south displacement field derived from the subpixel correlation of Sentinel-2 optical imagery. The yellow star indicates the location of the epicenter, and the green triangle marks the location of the footage, as given in the YouTube video description. The black lines show the surface trace of the finite fault model obtained from the USGS. Dots labeled A to F show the locations of the shallowest subfaults of the USGS finite fault model, which were used to extract the local slip-rate functions shown in (C). (**B**) Moment-rate functions from the USGS and SCARDEC. (**C**) Local slip-rate functions derived from the USGS finite fault model. This was only done for the shallowest subfaults of the segment highlighted in (A). Note that we use a triangular function to display the local slip-rate function instead of the asymmetric cosine function used by the USGS kinematic inversion.

slip duration is $\Delta t = 1.4$ s (Fig. 3A), independent of any assumptions related to spatial scaling. By contrast, our estimates of total and intermediate slip must consider a relative uncertainty of approximately ±20% in the spatial scaling factor. The uncertainty in total slip also affects absolute values of the slip rate.

Duration and shape of the slip-rate function

Our results highlight the much shorter local slip duration (1.4 s) compared with the total rupture duration of 100 to 120 s (Fig. 1B). This clearly shows that the rupture had a pulse-like nature, also called a self-healing pulse, at least at the surface along this segment of the fault. Self-healing pulses have long been identified as a possible mode of earthquake rupture (*14*). Indeed, pulse-like behavior appears to be a fairly common feature of rupture models for large earthquakes (*15*, *16*). The origin of this pulse-like behavior can be attributed to (i) the elongated geometry of the rupture or (ii) the structure of the fault damage zone (*17*). In the case of the Mandalay earthquake, the rupture was expected to have quickly saturated the seismogenic width, resulting in a large length-to-width ratio at the footage location. Under these conditions, it is expected that a pulse-like rupture would emerge naturally (*18*, *19*). However, alternative slip-healing mechanisms may also account for the short 1.4-s slip duration observed here. In particular, thermal pressurization of pore fluids within a thin gouge layer can induce rapid restrengthening behind the rupture tip. Using representative values for hydrothermal diffusivity ($\alpha = 10^{-4}$ to 10^{-6} m^2 s^{-1}) and a gouge thickness h ranging from 10^{-2} to 10^{-3} m (*20*, *21*), the expected slip duration scales to $\Delta t \approx h^2/\alpha \approx 1$ s (*22*), consistent with our observations. This mechanism does not rely on rupture aspect ratio and could explain the relatively small pulse width compared with the total rupture depth. Our results, however, are agnostic to various healing mechanisms simply because of the lack of additional in situ physical measurements.

The slip-rate values are consistent with the coseismic slip rates expected during large earthquakes. Particle velocity in the vicinity of the fault was predicted to be capable of reaching several meters per second for self-healing pulse ruptures (*14*). Note that our direct measurements of the fault slip rate are much higher than those estimated from inverted teleseismic data (*9*), which indicated a peak slip rate on the order of 0.2 m s^{-1} at the footage location (Fig. 1C). In addition, the recorded duration of local slip is much shorter than the inverted estimate: 1.4 s compared with ≈16 s for the USGS kinematic model. However, it is well known that modeling the rupture process involves substantial uncertainty (*6*, *23*). Indeed, the inverted slip duration just 10 km north of the footage location is much shorter (~6 s), and the peak slip rate is higher (0.45 m s^{-1}) (Fig. 1C). This discrepancy underscores the value of direct observation of slip during a natural earthquake as an empirical benchmark for validating and refining seismic source models of large earthquakes.

The shape of the slip-rate function is a key parameter in kinematic slip models (*24*) and plays an important role in ground motion prediction (*25*). In this study, the smoothed slip-rate function exhibits a simple form that can be described as an asymmetric triangle, with the duration of the acceleration phase being ~60% that of the deceleration phase. A similarly smooth, asymmetric, and triangular slip-rate function has been observed in friction experiments involving self-healing slip pulses propagating at 0.76 times the Rayleigh wave speed (*26*).

An analytical form of the slip-rate function was proposed for mode III self-healing pulses (*27*). This form, sometimes referred to as the Yoffe function, is characterized by an infinitely steep onset, a sharply peaked maximum, and a long tail with a variable slope. Although the Yoffe function qualitatively captures the asymmetry observed in our data, its overall shape is more peaked and strongly asymmetric. A regularized version of the Yoffe function, derived from dynamic rupture models (*28*), more closely resembles the slip-rate function observed in our measurements, though it still exhibits greater asymmetry.

Estimate of dynamic source parameters from the fault slip rate

We applied the elastodynamic equilibrium condition to infer the evolution of shear stress from a slip-based model. Assuming a two-dimensional plane strain shear (mode II) rupture propagating at a constant sub-Rayleigh rupture velocity V_r along the fault, the shear stress $\tau(x)$ is related to the spatial distribution of slip rate $v(x)$ through the singular integral equation (*22*, *29*, *30*)

$$\tau(x) = \tau_0 - \frac{\bar{\mu}}{2\pi V_r} PV \int_0^L \frac{v(\xi)}{x - \xi} d\xi \quad (1)$$

www.xsava.xyz

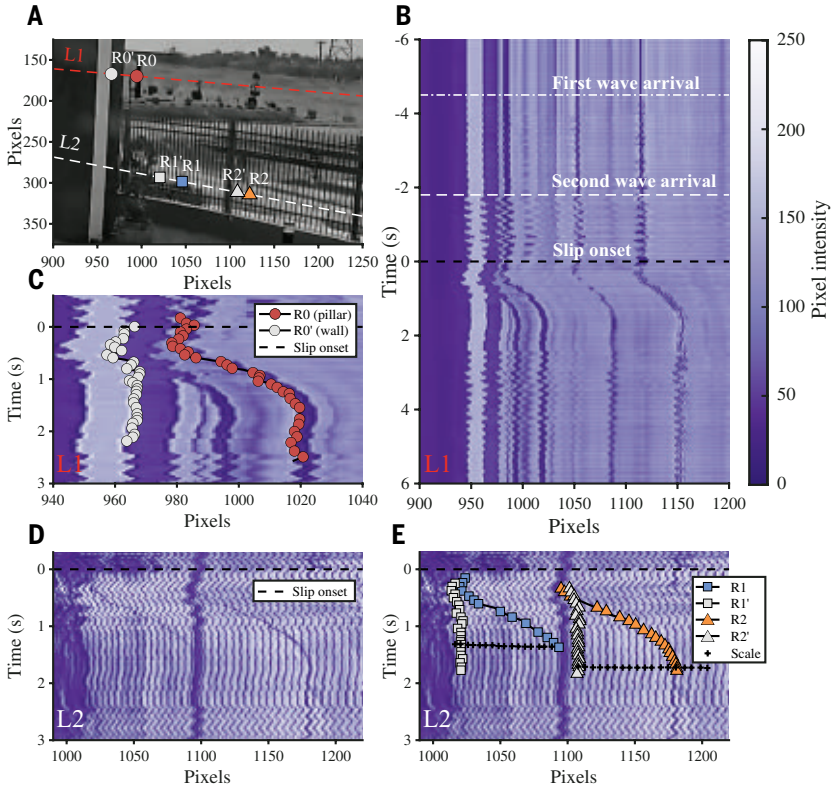


Fig. 2. Landmark tracking method. (**A**) Zoomed-in video frame showing tracked landmarks and their corresponding reference objects. The landmarks are as follows: *R0* (red circle) is the front-side pillar, *R1* (blue square) is the pole visible behind the fence, and *R2* (orange diamond) is the second pole behind the fence. The reference points are as follows: *R0′* (gray circle) is the arch wall, and *R1′* and *R2′* (gray square and diamond, respectively) are fence bars near the initial positions of *R1* and *R2*. The red dashed line (L1) indicates the axis used for projections in (B) and (C). The white dashed line (L2) indicates the axis used for projections in (D) and (E). (**B**) Motion tracking with annotations showing the arrival of the first seismic wave, second wave, and slip onset. (**C**) Pixel displacement time series for *R0* and *R0′*, measured along L1. (**D**) Pixel displacement time series for *R1*, *R1′*, *R2*, and *R2′*, measured along L2. (**E**) Same as (D), with additional annotations showing the tracking of each landmark and associated reference for *R1* and *R2* as well as a motion scale (black crosses).

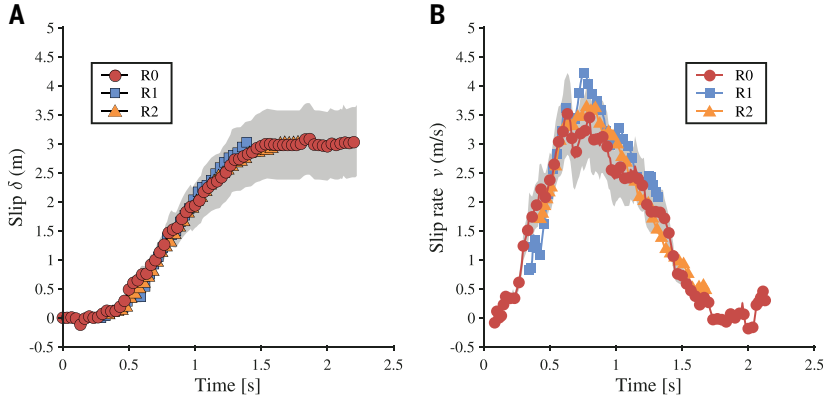


Fig. 3. Slip and slip-rate functions measured at the camera location. (**A**) Evolution of the fault slip during the rupture propagation modeled using the three landmark-reference pairs (Fig. 2A). (**B**) Slip-rate function derived from (A) for the three pairs. The gray shaded areas in (A) and (B) represent the effect of the 20% uncertainty in our slip estimate on the measurements derived from *R0*, which is the landmark used in the modeling. The symbols correspond to those described in Fig. 2.

where τ_0 is the background stress; $\bar{\mu}$ is a scaled shear modulus, which depends on the rupture velocity V_r; x is the distance to the rupture tip; $L = V_r \Delta t$ is the length of the slip pulse, which defines the spatial extent of the actively slipping region; ξ is the integration variable; and PV denotes the Cauchy principal value of the integral.

To compute the shear-stress evolution $\tau(x) - \tau_0$ along the fault, we interpolated $v(x)$ using a Chebyshev-based fit (Fig. 4, A and B), which captures the asymptotic behavior of slip and slip rate expected from fault mechanics, at both the leading and trailing edges of the pulse (see figs. S4 and S5 and supplementary materials for more details).

Once $v(x)$ was known, we inferred the shear stress evolution from Eq. 1 (Fig. 4C). Because the exact value of V_r was not known at this stage, we present results for two representative cases in Fig. 4C: $V_r/C_s = 0.75$ and $V_r/C_s = 0.90$. These ratios correspond to the two end-member rupture velocities of a rupture propagation scenario accounting for the ground motion and arrival times of the first and second seismic-wave arrivals as observed in the footage (Fig. 2B), implying a local sub-shear rupture velocity (see supplementary materials and fig. S3).

As anticipated for pulse-like ruptures, we observed that the stress drop was followed by a restrengthening phase (Fig. 4D) that spanned ~17% of the pulse length. The dependence of the shear-stress evolution on the rupture velocity V_r was also expected: A lower rupture velocity V_r requires a larger strength drop $\Delta\tau$, defined as the difference between the maximum and the minimum shear stress, to match the measured cumulative slip and slip rate. In addition, the size of the cohesive zone X_c, which corresponds to the distance over which the frictional strength of a fault degrades from its peak to residual level, increases linearly with V_r. Specifically, we found $\Delta\tau = 5.7$ MPa and $X_c = 3749$ m for $V_r = 0.90C_s$, compared with $\Delta\tau = 28.2$ MPa and $X_c = 3124$ m for $V_r = 0.75C_s$. Examining the apparent stress-slip relationship in Fig. 4D, we observed a decay of shear stress τ with slip δ, up to a critical slip value $D_c = 2.94$ m, which is consistent for both rupture velocities.

We then calculated the breakdown work $W_b = \int_0^{D_c} [\tau(\delta) - \tau_{min}] d\delta$ associated with the slip pulse, integrating up to the point of minimum stress τ_{min}, defined spatially at X_c and in cumulative slip at D_c (Fig. 4, C and D). Because this estimate is strongly dependent on final slip D_{fin} and rupture velocity V_r, we conducted a sensitivity analysis to characterize the uncertainty on these parameters (fig. S6 and supplementary materials). For the assumed final slip of 3.0 m, W_b ranges from 7.7 MJ m^{-2} for $V_r = 0.90C_s$ to 38.2 MJ m^{-2} for $V_r = 0.75C_s$.

To better constrain V_r and provide an independent direct measurement of the energy release rate G during the rupture pulse of the M_w 7.7 Mandalay earthquake, we fitted a classical two-dimensional steady-state rupture pulse model (*29*) to the observed slip pulse. The model assumes a steady-state sub-Rayleigh rupture velocity V_r and a constant-width cohesive zone in which the friction linearly decreases from the peak to the residual frictional strength ($f_p = 0.6$ and $f_r = 0.1$, respectively) behind the rupture front (fig. S7A). Unlike the direct slip-based model used previously, this model imposes a cohesive zone friction law and no healing within the pulse. Therefore, the energy released during

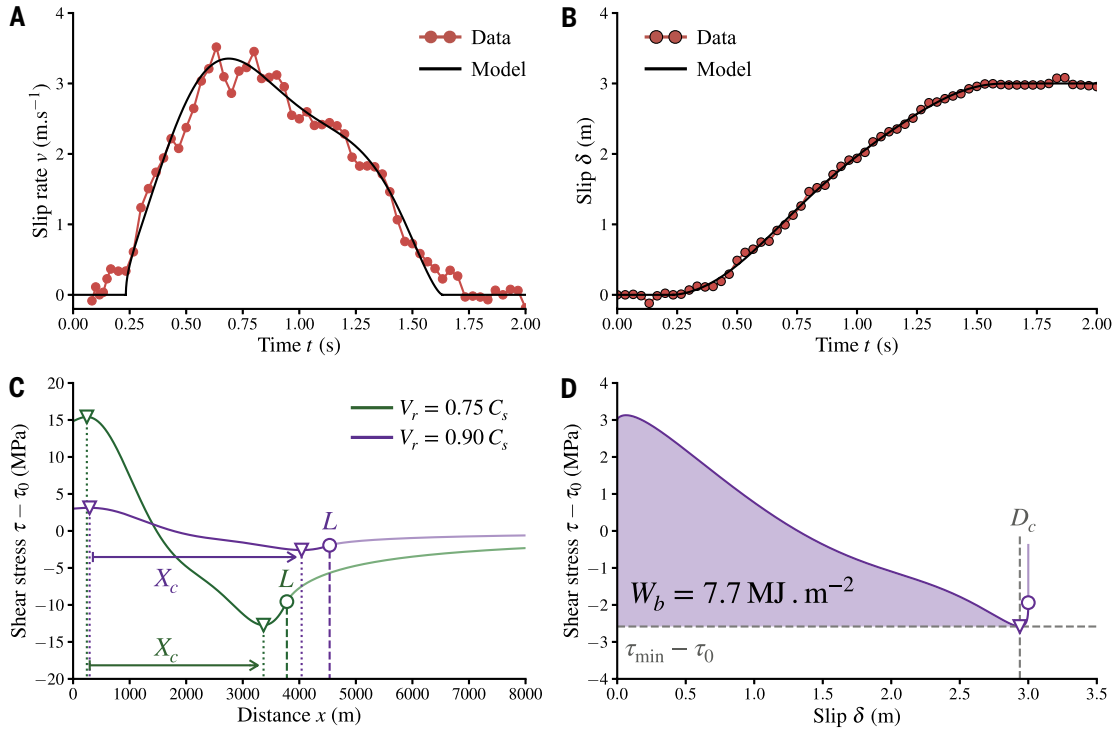


Fig. 4. Stress evolution and stress-slip relationship inferred from the slip-rate function. (**A** and **B**) Direct measurement (red circles) and Chebyshev-based fit (solid black curve) of (A) fault slip rate and (B) fault slip at the footage location. (**C**) Solution of the evolution of shear-stress change as a function of the distance from the rupture tip during slip-pulse propagation for rupture scenarios $V_r = 0.75C_s$ (solid green curve) and $V_r = 0.90C_s$ (solid purple curve). Cohesive zone size X_c is defined as the distance between the maximum and minimum of the shear-stress change (triangles), whereas the pulse length L corresponds to the extent of the actively sliding region (circles). (**D**) Evolution of stress variation $\tau - \tau_0$ with slip, including the healing stage for $V_r = 0.90C_s$. The breakdown work W_b (in shaded purple area) is defined from an integral of the shear-stress change $\tau - \tau_{min}$ up to the critical slip distance $\delta = D_c$.

the weakening contributes entirely to the rupture propagation and corresponds to G as described in fracture mechanics. Whereas the direct slip-based approach is applicable in this study because of the availability of true on-fault slip measurements from CCTV footage, the linear cohesive zone model is a more general framework that can be applied to conventional seismic observations, including data recorded at stations located away from the fault (*31*). We assumed a nominal normal traction of $\sigma_0 = 10$ MPa, consistent with the shallow nature of the rupture. The slip pulse duration was fixed at $\Delta t = 1.4$ s, as estimated previously. We performed a grid search on the scaled cohesive zone size to X_c/L and scaled rupture velocities V_r/C_s, identifying the parameters that produce the largest signal similarity index (SSI), which is a Euclidean-norm measure of how well the predicted slip-rate evolution matches that measured from image analysis. As seen in Fig. 5A, the optimal SSI corresponds to $X_c/L = 0.71$ and $V_r/C_s = 0.903$, which provides the fit of the curve in Fig. 5B. This X_c/L value is lower than the elastodynamic estimate of 0.83, as the pulse model prohibits restrengthening within the pulse, reducing the effective weakening (cohesive) zone. Nevertheless, the result supports $V_r = 0.90C_s = 3240$ m s^{-1} as the most consistent rupture scenario.

We have now obtained a complete mechanical characterization of the slip pulse observed in the CCTV data. The inferred strength drop was $\tau_p - \tau_r = 5$ MPa, the stress drop was $\tau_0 - \tau_r = 2$ MPa, and the corresponding slip-weakening distance was $X_c \approx 2.5$ m. These values are consistent with the slip-based model estimates ($\tau_p - \tau_r = 5.7$ MPa, $\tau_0 - \tau_r = 2.6$ MPa), except for $D_c = 2.94$ m, which was slightly overestimated because of the larger cohesive zone size. Ultimately, the energy release rate G was computed as $\int_0^{D_c} [\tau_f(\delta) - \tau_r] d\delta = 5.8$ MJ/m^2, where $\tau_f(\delta)$ is the strength evolution with slip δ (Fig. 5C). Assuming that $G = G_c$, our analysis here provides a direct estimate of the local fracture energy G_c of a natural fault. The energy release rate is of the same order of magnitude as our elastodynamic estimate of the breakdown work $W_b = 7.7$ MJ m^{-2}. Although related, the two quantities are not equivalent: G_c is expected to be smaller than W_b for ruptures that have undergone complete frictional weakening (*32*–*34*). Our value of G_c is consistent with expectations for an earthquake of the size and geometry of the Mandalay event (*35*–*38*), but it is substantially larger than values typically reported for experimental faults, both in analog materials (*39*, *40*) and in crustal rock samples (*41*, *42*). This difference is primarily attributed to the larger slip-weakening distance D_c inferred here, compared with the much smaller values typically observed during laboratory earthquakes (*40*, *41*). Nevertheless, such large values of D_c remain compatible with those measured during high-velocity friction experiments under upper-surface stress conditions (*43*–*45*) and are expected from seismic kinematics inversion of large earthquakes (*46*). Beyond the specific case of the Mandalay earthquake, these direct measurements offer a rare, empirical benchmark for earthquake source modeling. Unlike traditional kinematic and dynamic inversions, which are limited by data coverage and modeling assumptions, the video-derived slip and slip-rate functions provide ground-truth constraints on the physical processes behind large earthquakes, such as fault weakening, energy dissipation, and rupture dynamics.

Discussion

Our study highlights the potential for a new observational approach in earthquake science: strategically deploying CCTV or high–frame rate cameras near active, shallow-fault zones. Such installations could capture not only rapid, coseismic fault slip, as demonstrated here, but also slower aseismic deformation or postseismic creep. Compared with

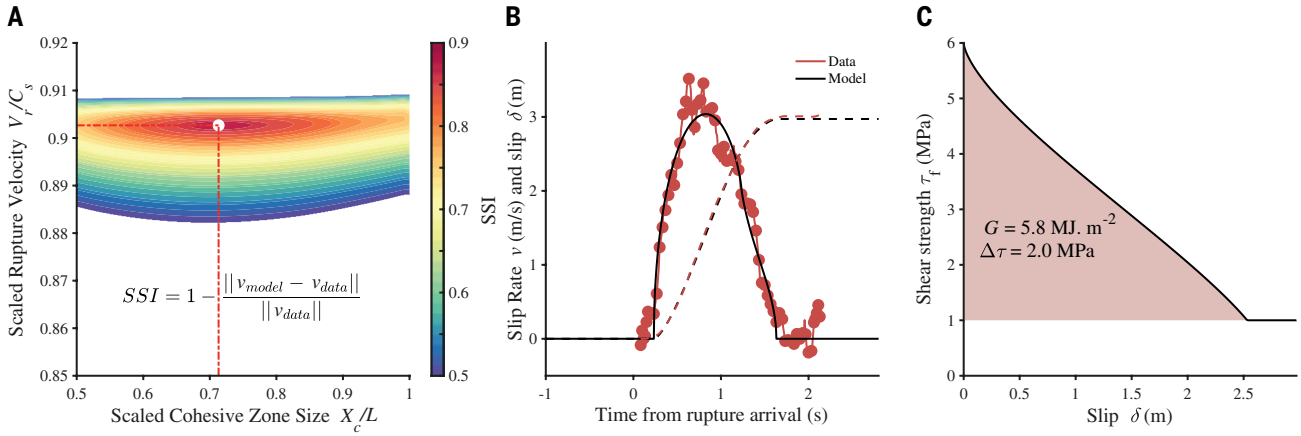


Fig. 5. Evaluation of rupture velocity and energy release rate. (**A**) Grid search for the best-fitting scaled cohesive zone size, X_c/L, and scaled rupture velocity, V_r/C_s. The fit is measured using the SSI, as defined in the text. v_{model} and v_{data} are slip rates from the model and the data, respectively. (**B**) Comparison of slip (dashed lines) and slip rate (solid lines) between the observed and best-fitting model slip pulse. (**C**) Evolution of friction strength, τ_f, with slip, δ.

traditional instrumentation such as broadband seismological stations or continuous GPS networks, video monitoring systems are relatively low cost, widely available, and easy to deploy, making them an attractive complementary tool for expanding fault-monitoring capabilities. With appropriate placement and calibration, these instruments could provide direct, high-resolution constraints on fault behavior across the full spectrum of slip modes, opening new avenues for studying earthquake physics.

REFERENCES AND NOTES

1. J.-P. Avouac, *Annu. Rev. Earth Planet. Sci.* **43**, 233–271 (2015).
2. A. Tarantola, B. Valette, *Rev. Geophys.* **20**, 219–232 (1982).
3. A. Saraò, S. Das, P. Suhadolc, *J. Seismol.* **2**, 1–25 (1998).
4. I. A. Beresnev, *Bull. Seismol. Soc. Am.* **93**, 2445–2458 (2003).
5. S. Hartzell, P. Liu, C. Mendoza, C. Ji, K. M. Larson, *Bull. Seismol. Soc. Am.* **97**, 1911–1934 (2007).
6. P. M. Mai *et al.*, *Seismol. Res. Lett.* **87**, 690–708 (2016).
7. G. Bertrand, C. Rangin, *J. Asian Earth Sci.* **21**, 1139–1157 (2003).
8. N. G. Reitman *et al.*, Remote surface rupture observations for the M7.7 2025 Mandalay, Burma (Myanmar) earthquake. USGS Data Release (2025); https://doi.org/10.5066/P1RYMWCK.
9. US Geological Survey, Finite Fault – M 7.7 - 2025 Mandalay, Burma (Myanmar) earthquake (2025); https://earthquake.usgs.gov/earthquakes/eventpage/us7000pn9s/finite-fault [accessed 16 May 2025].
10. M. Vallée, J. Charléty, A. M. Ferreira, B. Delouis, J. Vergoz, *Geophys. J. Int.* **184**, 338–358 (2011).
11. N. Inoue *et al.*, *Seismica* **4**, 1 (2025).
12. H. Richter, "Watch an earthquake split a hillside in two," *Science*, 22 July 2025.
13. J. Kearse, Y. Kaneko, *Seismic Record* **5**, 281–288 (2025).
14. T. H. Heaton, *Phys. Earth Planet. Inter.* **64**, 1–20 (1990).
15. D. Melgar, G. P. Hayes, *Geophys. Res. Lett.* **44**, 9691–9698 (2017).
16. V. Lambert, N. Lapusta, S. Perry, *Nature* **591**, 252–258 (2021).
17. F. Barras, E. Aharonov, F. Renard, *Geophys. Res. Lett.* **52**, e2024GL111189 (2025).
18. H. Weng, J.-P. Ampuero, *J. Geophys. Res. Solid Earth* **124**, 8584–8610 (2019).
19. H. Weng, J.-P. Ampuero, *Nat. Commun.* **13**, 7327 (2022).
20. J. R. Rice, *J. Geophys. Res. Solid Earth* **111**, B05311 (2006).
21. A.-A. Gabriel, D. I. Garagash, K. H. Palgunadi, P. M. Mai, *Science* **385**, eadj9587 (2024).
22. D. I. Garagash, *J. Geophys. Res. Solid Earth* **117**, B04314 (2012).
23. S. Minson, M. Simons, J. Beck, *Geophys. J. Int.* **194**, 1701–1726 (2013).
24. H. N. Razafindrakoto, P. M. Mai, *Bull. Seismol. Soc. Am.* **104**, 855–874 (2014).
25. I. A. Beresnev, *Pure Appl. Geophys.* **181**, 1859–1869 (2024).
26. G. Lykotrafitis, A. J. Rosakis, G. Ravichandran, *Science* **313**, 1765–1768 (2006).
27. S. Nielsen, R. Madariaga, *Bull. Seismol. Soc. Am.* **93**, 2375–2388 (2003).
28. E. Tinti, E. Fukuyama, A. Piatanesi, M. Cocco, *Bull. Seismol. Soc. Am.* **95**, 1211–1223 (2005).
29. J. R. Rice, C. G. Sammis, R. Parsons, *Bull. Seismol. Soc. Am.* **95**, 109–134 (2005).
30. N. Brantut, D. I. Garagash, H. Noda, *J. Geophys. Res. Solid Earth* **124**, 8998–9020 (2019).
31. E. M. Dunham, R. J. Archuleta, *Bull. Seismol. Soc. Am.* **94** , S256–S268 (2004).
32. F. Paglialunga, F. Passelegue, M. Lebihain, M. Violay, *Earth Planet. Sci. Lett.* **626**, 118550 (2024).
33. D. S. Kammer *et al.*, *Nat. Commun.* **15**, 4736 (2024).
34. B. Fryer, M. Lebihain, C. Noel, F. Paglialunga, F. Passelègue, *J. Mech. Phys. Solids* **193**, 105876 (2024).
35. R. E. Abercrombie, J. R. Rice, *Geophys. J. Int.* **162**, 406–424 (2005).
36. E. Tinti, P. Spudich, M. Cocco, *J. Geophys. Res. Solid Earth* **110**, B12303 (2005).
37. S. Nielsen *et al.*, *J. Seismol.* **20**, 1187–1205 (2016).
38. M. Cocco *et al.*, *Annu. Rev. Earth Planet. Sci.* **51**, 217–252 (2023).
39. I. Svetlizky, J. Fineberg, *Nature* **509**, 205–208 (2014).
40. F. Paglialunga *et al.*, *Earth Planet. Sci. Lett.* **584**, 117442 (2022).
41. D. S. Kammer, G. C. McLaskey, *Earth Planet. Sci. Lett.* **511**, 36–43 (2019).
42. S. Xu, E. Fukuyama, F. Yamashita, *J. Geophys. Res. Solid Earth* **124**, 766–787 (2019).
43. A. Tsutsumi, T. Shimamoto, *Geophys. Res. Lett.* **24**, 699–702 (1997).
44. T. Hirose, T. Shimamoto, *J. Geophys. Res. Solid Earth* **110**, B05202 (2005).
45. G. Di Toro *et al.*, *Nature* **471**, 494–498 (2011).
46. T. Mikumo, K. B. Olsen, E. Fukuyama, Y. Yagi, *Bull. Seismol. Soc. Am.* **93**, 264–282 (2003).
47. F. Passelègue *et al.*, Data and model for Direct Estimation of Earthquake Source Properties from a Single CCTV Camera. Zenodo (2025); https://doi.org/10.5281/zenodo.16785672.

ACKNOWLEDGMENTS

We thank those who uploaded and shared the analyzed footage on social media. F.P. thanks B. Fryer, N. Calza, and F. Paglialunga for discussions and for proofreading the manuscript. S.L. and F.P. thank J.-P. Ampuero and F. Courboulex for valuable discussions. We thank D. Garagash and an anonymous reviewer for their constructive remarks, which helped to enhance this paper. **Funding:** F.P. acknowledges support from the European Union (ERC Starting Grant HOPE, no. 101041966). H.S.B. acknowledges support from ERC Consolidator Grant PERSISMO (no. 865411) for partial support of this work. Q.B. acknowledges support from ERC Starting Grant EARLI (no. 949221) for partial support of this work. **Author contributions:** S.L. and F.P. conceived the study; S.L. and K.W.H. conducted the frame analysis based on satellite imaging; M.L. developed the elastodynamic inversion; F.P., H.S.B., and M.L. estimated the rupture properties; C.T. and Q.B. analyzed the kinematic inversion from USGS and provided estimates of the wave speeds; S.L. and F.P. wrote the first draft of the manuscript; and K.W.H. provided critical data for spatial scaling and validation. All authors contributed to discussions, interpretation of results, and review of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data and code necessary to directly reproduce our results are available online at Zenodo (*47*). **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. This research was funded in whole or in part by ERC (101041966, 865411, and 949221); as required the author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adz1705
Supplementary Text; Figs. S1 to S7; References (*48–54*)

Submitted 20 May 2025; accepted 29 August 2025

10.1126/science.adz1705

RESEARCH ARTICLE

Supershear rupture sustained through a thick fault zone in the 2025 M_w 7.8 Mandalay earthquake

Shengji Wei[1,2,3]*†, Xin Wang[4,5]†, Chenglong Li[6], Hongyu Zeng[2], Zhangfeng Ma[2], Qibin Shi[7], Han Chen[3,6], Yihe Huang[8], Mingzhe Lyu[2,3], Jiangtao Liao[3,9], Shun Yang[10], Yiming Bai[11], Phyo Maung Maung[2], Kyawmoe Oo[12], Yin Myo Min Htwe[12], Jianyong Zhang[13], Luca Dal Zilio[2,3], Xinjian Shan[5], Ling Chen[14,5]

On 28 March 2025, a moment magnitude (M_w) 7.8 earthquake ruptured a seismic gap along the Sagaing fault in Myanmar, generating a surface rupture 480 km in length and causing widespread damage. With holistic geodetic and seismic techniques, we resolved its rupture dynamics and fault zone structure. The rupture initiated as bilateral subshear and transitioned to supershear (~5.3 km/s) about 100 km south of the epicenter, sustaining this velocity for more than 200 km. The supershear segment aligns with a ~2-km–thick low-velocity fault zone exhibiting ~45% shear wave speed reduction. We suggest that the thick fault zone, aided by fault geometry and basin structure, enabled prolonged supershear propagation. Our findings emphasize that observations and models encompassing the entire fault zone are crucial for advancing both specific-event and earthquake-cycle simulations.

The largest continental crustal earthquakes—capable of producing ruptures stretching hundreds of kilometers and affecting multiple populated centers—pose considerable seismic threats. Such events often develop into supershear ruptures in which the rupture speed exceeds crustal shear wave velocity, as reported for recent strike-slip events such as the 2001 M_w 7.8 Kokoxili earthquake (*1*–*3*), the 2002 M_w 7.9 Denali earthquake (*3*–*5*), and the 2023 M_w 7.8 Turkey earthquake (*5*–*8*). Notably, the supershear rupture portion of these events often take place on relatively simple fault geometries (*1*, *9*, *10*) or "superhighways" (*11*) and therefore have reduced energy dissipation into secondary fault branches, allowing the rupture to develop to supershear speed and form long surface ruptures (*9*). Continental strike-slip faults at plate boundaries, such as the North Anatolian (Türkiye), Sagaing (Myanmar), and San Andreas (USA) faults, are particularly concerning due to their relatively simple geometries and historical occurrence of large events.

During an earthquake rupture, part of the released energy contributes to creating damage zones or/and fault zones, characterized by diffuse inelastic coseismic deformation or substantial reduction of seismic wave speeds, as increasingly reported for active continental faults (*12*–*19*). Here, we term them coseismic damage zones when created during earthquake ruptures, primarily constrained by near-fault geodetic and field deformation observations (*15*, *17*, *20*, *21*), or low-velocity fault zones (LVFZ) when indicating the cumulative effects of historical ruptures (*19*, *22*–*26*).

However, clarity is still lacking regarding the connection between these fault zones and large strike-slip ruptures, and the role of fault zones in long, sustained supershear ruptures. The recent M_w 7.8 earthquake on 28 March, 2025 in Mandalay, Myanmar, provides a valuable case to address these questions. The event took place on the Sagaing fault, which forms the eastern boundary of the Burma microplate, a sliver block created by oblique collision between the Indian Ocean and the Sunda plates (*27*) (Fig. 1A). The fault accommodates a tectonic slip rate of 20 to 23 mm/yr (*28*) (Fig. 1B). Over the past 200 years, the Sagaing fault has hosted more than 10 M > 6.5 earthquakes. However, a ~250-km seismic gap between the cities of Mandalay and Naypyidaw had remained unruptured since at least 1897—until the 2025 M_w 7.8 event (Fig. 1C) (*29*). In Myanmar, the 2025 event has so far directly resulted in at least 5352 fatalities, 7108 injuries, and 538 missing individuals (*30*). The impacts extended beyond Myanmar, including a tall building collapse in Bangkok, Thailand, located ~1000 km from the epicenter (*31*), and substantial building damage in Yunnan province, China (*32*). Early reports showed that the event produced a >450-km surface rupture and supershear rupture speed (*33*, *34*) and also triggered seismicity in Shan-Thai plateau and southern China (*35*).

Through the application of multiple techniques to both seismic and geodetic datasets, we investigated the events of its surface deformation, resulting coseismic damage zone, detailed rupture process, and the associated fault zone structures. These results offer new insights into the interrelationship between supershear rupture and fault zone structure.

3D surface deformation and damage zone

To characterize surface deformation and the ruptured fault associated with the earthquake, we first analyzed multiple sets of satellite observations. We employed three independent remote sensing techniques: differential interferometric synthetic aperture radar, aperture interferometry, and pixel offset on Sentinel-1/2 and ALOS-2 satellite imagery, to obtain fourteen coseismic surface displacement images (figs. S1 and S2 and methods in the supplementary materials). These datasets, acquired from multiple viewing geometries, provide complementary sensitivity to different components of ground deformation, with an effective spatial resolution of ~200 m. We integrated these observations to reconstruct the full three-dimensional (3D) coseismic displacement field produced by the earthquake (fig. S3). The results reveal a rupture length of ~480 km, characterized by predominant right-lateral strike-slip motion (Fig. 2 and fig. S3). Vertical displacements are minor (Fig. 2C), indicating that the rupture was nearly pure strike-slip.

The coseismic surface displacements exhibit clear along-strike variations. To the first order, the north–south component of displacement (Fig. 2A) reveals a three-segment pattern. The northern segment, located north of Mandalay, shows the largest displacement for the event—up to ~5.5 m. Deformation was greater to the east of the fault, suggesting an east-dipping fault geometry. The central segment, extending approximately 80 km southward from Sagaing, had the smallest displacement magnitude among the three segments and near-symmetric deformation across the fault. The southern segment shows larger displacement than the middle segment but the amplitude on the western side of the fault gradually increases compared with the eastern side, implying a transition to a west-dipping fault geometry.

Fault slip and coseismic damage-zone width derived from high-resolution (10-m) Sentinel-2 displacements also reveal distinct along-strike variations (Fig. 3A and methods). North of the epicenter, sharp

[1]Key Laboratory of Deep Petroleum Intelligent Exploration and Development, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China. [2]Earth Observatory of Singapore, Nanyang Technological University, Singapore. [3]Asian School of the Environment, Nanyang Technological University, Singapore. [4] Key Laboratory of Planetary Science and Frontier Technology, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China. [5]College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, China. [6]State Key Laboratory of Earthquake Dynamics and Forecasting, Institute of Geology, China Earthquake Administration, Beijing, China. [7]Department of Environmental and Planetary Sciences, Rice University, Houston, TX, USA. [8]Department of Earth and Environment Sciences, University of Michigan, Ann Arbor, MI, USA. [9]Key Laboratory of Earth Exploration and Information Technology of Ministry of Education, Chengdu University of Technology, Chengdu, China. [10]Marine Science and Technology College, Zhejiang Ocean University, Zhoushan, China. [11]School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore. [12]Department of Meteorology and Hydrology, Naypyitaw, Myanmar. [13]China Earthquake Networks Center, Beijing, China. [14]State Key Laboratory of Lithospheric and Environmental Coevolution, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China.
***Corresponding author. Email: shjwei@gmail.com** †These authors contributed equally to this work.

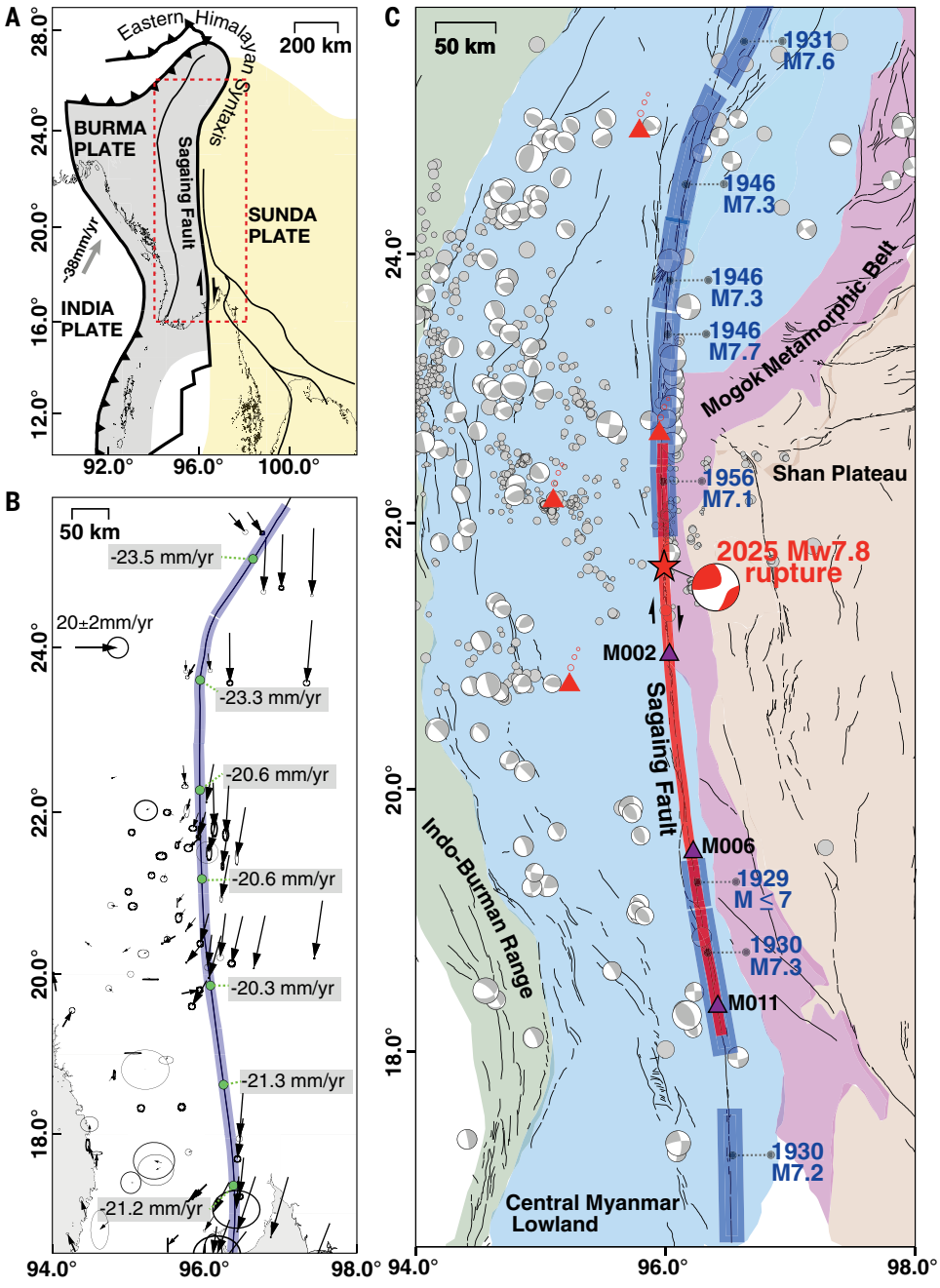


Fig. 1. Seismotectonic background. (**A**) Tectonic setting for the study region. The red dashed rectangle marks the locations of (**B** and **C**). (B) Slip rate distribution along the Sagaing Fault along with the GNSS observations used to derive these rates (*28*). GNSS velocities are relative to the Burma plate. (C) Historical earthquakes in the region. Gray beach balls represent focal mechanisms from the global centroid moment tensor (GCMT) catalog and (*60*). Gray dots are earthquake epicenters from the USGS catalog. Red triangles indicate locations of Quaternary volcanoes (*61*). The detailed fault inventory map is from (*27*) and the dark red line is our mapped ruptured fault trace of the 2025 event. The red star is our refined epicenter and the red beach ball is the GCMT solution.

fault-parallel offsets suggest a highly localized rupture zone with a narrow damage zone. South of the epicenter, the gradient progressively broadens across the fault, marking a transition to a wider damage zone (fig. S4). The measured fault total slip is ~1 to 6 m (Fig. 3C), distributed within a coseismic damage zone ~20 to 800 m wide (Fig. 3D). Along the fault segment north of the epicenter, where maximum coseismic slip occurred, the narrowest damage zone is ~60 m wide (Fig. 3D and fig. S4). The coseismic damage zone to the south broadens to between 90- and 190-m wide and can be divided into three sections by thickness (Fig. 3D): ~120 m from ~20 km north to ~210 km south of the epicenter, ~90 m from ~210 to 330 km south, and ~190 m further south.

Overall, the average width of these coseismic damage zones is ~100 m (Fig. 3D), which lies at the lower end of the observed range compared with other continental strike-slip earthquakes (*15*, *17*, *18*, *36*). This is due to the relatively simple geometry of the Sagaing fault, characterized by considerably fewer stepovers, bends, and branches. The segmented variation in damage zone width is more pronounced than the total slip. This contrast suggests that coseismic damage width is more strongly influenced by other factors, such as rock properties, compared with slip amplitude (*18*, *21*). Indeed, south of the epicenter, sediment thickness increases and near-surface shear wave velocities (top 1 km) decrease (Fig. 3, D and E), as revealed by detailed shallow velocity structure imaged along the Sagaing fault (*37*).

Static slip model

We constrained the fault geometry and coseismic slip distribution by inverting the 3D surface deformation fields (see methods). Our preferred model resolved nine fault segments along the ~480-km rupture, achieving a variance reduction of 99.6% (figs. S5 and S6). The simple fault geometry—requiring no bifurcations or branches (black line in Fig. 3A)—contrasts with the fault branching mechanisms proposed for large strike-slip events (*38*). This simple geometry also deviates from the empirical relationship between rupture length and number of fault segments, as summarized for other events (*39*, *40*), although the number of asperities may match the relationship better. Based on (*39*), a continental strike-slip event with a 480-km rupture length should correspond to 27 fault segments, much more than that observed for the 2025 event. Modest dip variations are evident along the strike, with the fault dipping ~70° eastward in the north, near-vertical in the central section, and slightly (~80°) west-dipping in the south (fig. S5).

The static slip model reveals that most coseismic slip is concentrated shallower than 12 km (Fig. 3B), consistent with the depth extent of interseismic locking inferred from global navigation satellite system (GNSS) observations (*28*). The slip distribution exhibits pronounced along-strike heterogeneity. The longest, also the largest, slip patch is located north of the epicenter and extended more than ~100 km with a peak slip of ~6 m. Southward, slip decreased to ~2.5 to 3 m along a ~70-km section, then increased again to ~4 m and remained relatively uniform over the remaining ~230 km of rupture. Superimposed on this broad pattern are short-wavelength (~10 to 20 km) variations, which are also independently detected in fault-crossing offsets from Sentinel-2 optical imagery.

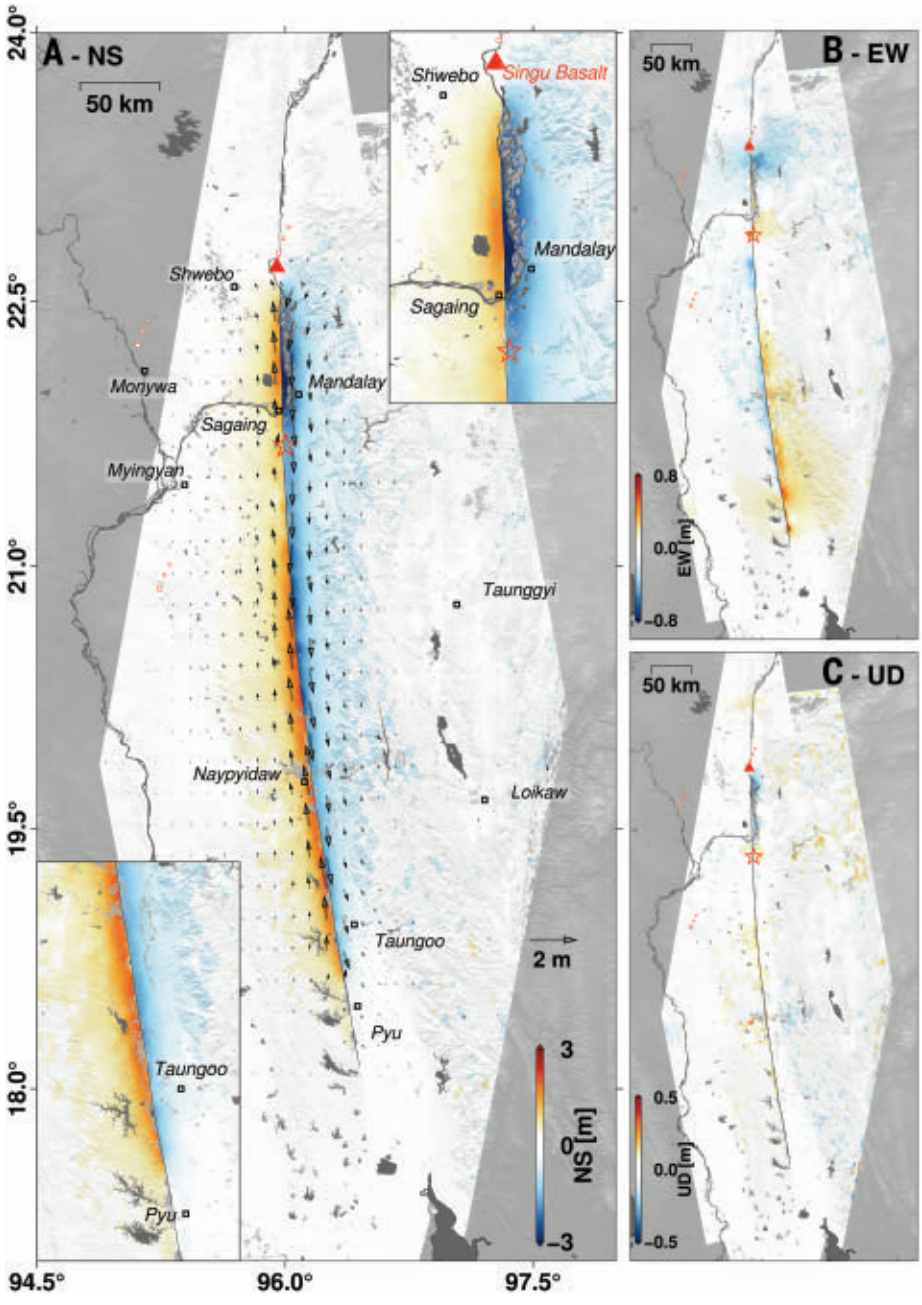


Fig. 2. 3D coseismic surface deformation map derived from satellite observations. (**A**) North-South direction deformation; (**B**) East-West; (**C**) up-down. The vectors in (A) represent horizontal motion and the black line is our mapped fault trace. Our refined epicenter is indicated by the red star. ALOS2 and Sentinal-1/2 images acquired from February to April 2025 were used to derive these 3D deformation maps (*62*) (see tables S1 and S2 for more details).

The measured total slip closely matches the near-surface slip estimated by our static slip model (Fig. 3, A and C). Checkerboard resolution tests confirmed the model's ability to resolve slip asperities across a range of spatial wavelengths (fig. S7). The averaged static stress drop estimated from our slip model was ~3 megapascals (MPa), which is consistent with that inferred for other major strike-slip faults (*41*).

Distinct rupture terminations were identified at both ends of the fault. To the north, the rupture terminated abruptly south of the Singu Basalt (*42*), coinciding with the southern limit of the 2012 M_w 6.8 Thabeikkyin earthquake rupture (Fig. 3B and fig. S8). We infer that a strong stress shadow posed by the 2012 event, combined with the mechanical contrast between basalt and the surrounding crust, formed an effective barrier to further rupture propagation. Toward the southern end, slip amplitude and rupture depth gradually diminished, suggesting a more distributed termination process. This segment overlaps the rupture zone of the 1929 and 1930 earthquakes (Fig. 3B), although the absence of detailed historical slip models precludes further constraints. These observations explain the displacement contrast patterns shown at the northern and southern termini, shown in the insets of Fig. 2A.

Shallow slip deficit (SSD) across the rupture was overall negligible, with slip on the near-surface subfault (~1 km depth) exhibiting only ~1% deficit relative to peak slip at depth (fig. S9). This value is substantially lower than the SSD observed in previous large strike-slip earthquakes (*43*–*45*), but closer to that reported for more recent events (*16*). The negligible SSD is robustly resolved by the dense and near-fault geodetic coverage for the 2025 event, as also suggested by recent studies of other earthquakes (*44*, *46*, *47*). However, along-strike variations occurred: On the fault segment south of Mandalay, where the coseismic slip amplitude was less, SSD (~7%) was larger than along the other fault segments (near 0%) (fig. S9). Reanalysis of interseismic GNSS velocity fields reveals that this segment is partially creeping, with a shallow creep rate of approximately 2.3 mm/yr extending to a depth of ~9.5 km (fig. S10), in contrast to the nearly fully locked behavior of adjacent fault segments. This contrast between SSD and coseismic slip suggests that depth-dependent and laterally varying fault frictional properties contributed to observed coseismic slip variations.

Kinematic and supershear rupture

We characterized the first-order rupture evolution using a path-calibrated teleseismic back-projection (BP) (*5*) technique on high-frequency [(HF), 0.3 to 1.0 Hz, fig. S11] seismic data from the Australian (AU) and European (EU) arrays (fig. S12A and methods). Locations of the HF radiators from BP align closely with the fault trace (Fig. 4A). The size of the BP radiators is proportional to the power of the stacked teleseismic waveforms. Our result shows that the rupture initially propagated bilaterally, with northward and southward segments each rupturing ~100 km at subshear speeds (~3 km/s). The southward rupture then accelerated to a supershear speed of ~5.3 km/s, maintaining this speed for at least 200 km (Fig. 4C). The robustness of the result is verified by BP applications to smaller nearby events consistent with their locations in a refined catalog (fig. S12). Synthetic tests mimicking mainshock rupture further confirmed the robustness of the result (figs. S13 and S14), which also shows that the array located away from the rupture directivity had less sensitivity (fig. S13, methods).

The evolution of rupture speed is further supported by the location and timing of subevents from a multiple point source (*48*) (MPS)

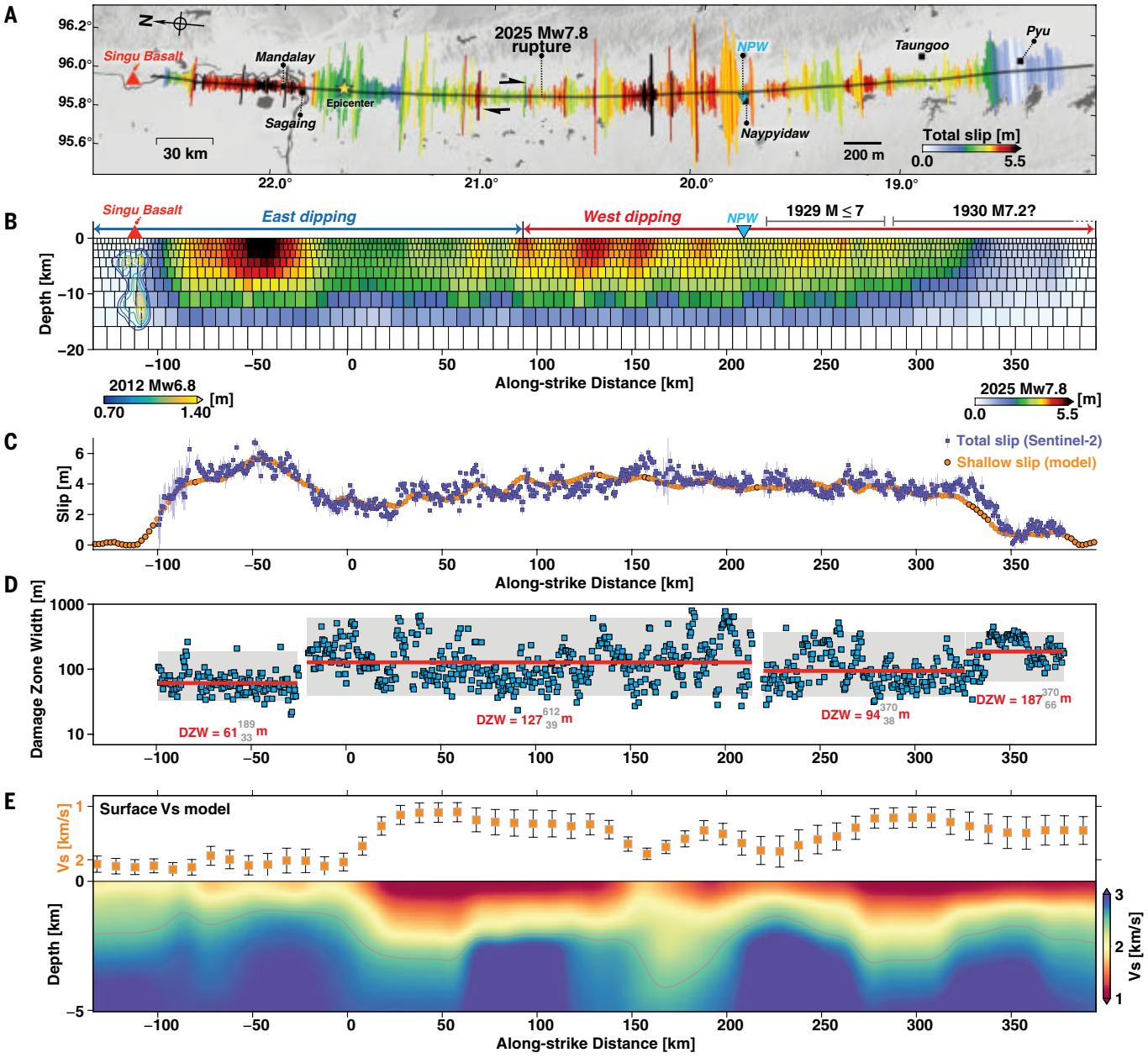


Fig. 3. Coseismic damage zone width and slip model, in comparison with velocity model. (**A**) Coseismic damage zone width and surface slip (*62*) estimated from Sentinel-2 image correlation. The length of the fault-perpendicular lines represents the width of the coseismic damage zone and the colors indicate the total slip magnitude. (**B**) Depth view of geodetic slip distribution for the 2025 M_w 7.8 event and the slip model of the 2012 M_w 6.8 event are shown in color-coded contours. Rupture extent for the 1929 and 1930 earthquakes is indicated at the southern end of the rupture. The purple triangle indicates the location of the NPW strong motion station. (**C**) Comparison between fault total slip from Sentinel-2 displacement [purple squares with uncertainty defined as standard deviation (one-sigma) between fault-parallel displacement (10 m binned mean) and fitting displacement] and shallow (≤1 km depth) slip amplitudes in the static slip model (orange dots). (**D**) Damage zone width (blue squares). The red lines are the averaged damage zone width of four segments. (**E**) Shear wave speed model in the top 5 km from (*37*). Orange squares represent the numerical average of uppermost 1-km surface shear wave velocity and uncertainty, in which the uncertainty is defined in (*37*). The gray line in (E) indicates the sedimentary base.

inversion using local-to-regional strong motion waveform and teleseismic body waves (Fig. 4B, fig. S16, and methods). MPS results fit the observations closely, with an average waveform cross-correlation coefficient of 0.8 across all components and stations (fig. S17). The result consists of 10 subevents (fig. S18), and the spatiotemporal evolution of these subevents shows that the mainshock began with a bilateral rupture that transitioned to supershear propagation ~100 km south of the epicenter (Fig. 4C). The overall source time function captures the low-frequency component of the moment-rate function derived from the finite fault inversion (Fig. 4D, inset; see details below).

Strong-motion waveforms recorded at the Naypyidaw (NPW) station (*49*), located 2.6 km from the fault trace at 19.779°N (blue triangle in Fig. 4A), provide independent evidence for supershear rupture propagation. To ensure the reliability of the absolute timing and waveform data from this station, we verified the accuracy of its GNSS time synchronization by comparing teleseismic arrival times with those recorded globally (fig. S15 and methods).

We then refined the mainshock epicenter and origin time using a nearby, well-located calibration event (fig. S19 and methods). Our analysis yields a revised epicenter at the latitude of 21.641°N, ~37 km

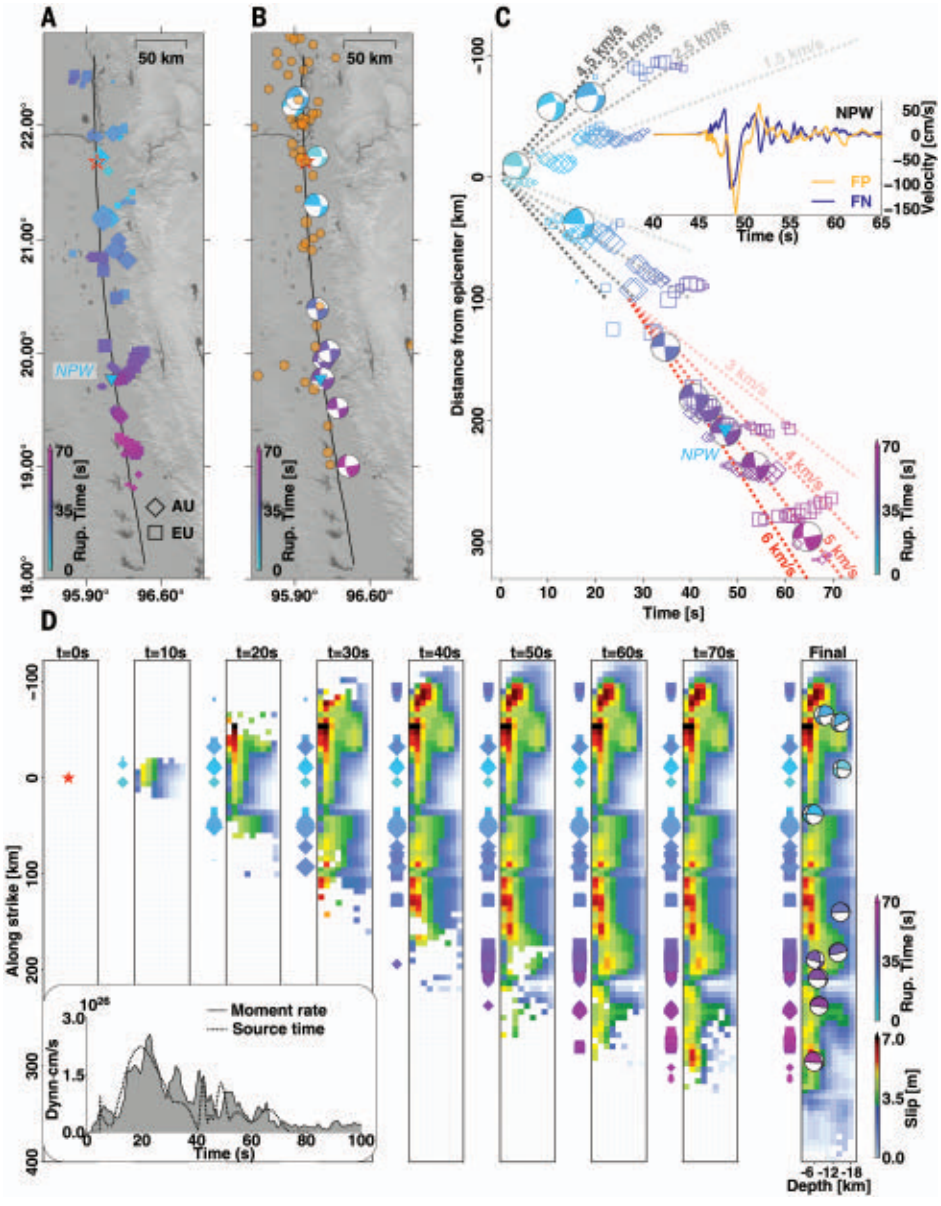


Fig. 4. Kinematic rupture models. (**A**) Map view of the back-projection (BP) results (*62*) from the EU array (squares) and the AU array (diamonds), color coded by time and sized by the power of stacked teleseismic waveforms. (**B**) MPS results (*62*) in map view, in which beach balls are color-coded by peak time in the source time function, projected to the low hemisphere. Orange dots indicate aftershocks provided by the USGS catalog. The blue triangles in (A) and (B) represent the location of the NPW station. (**C**) Rupture speed analysis and horizontal component velocity waveform at the NPW station. Fault-parallel (FP) and fault-normal (FN) directions are relative to the fault azimuth angle of 173 degrees. (**D**) Finite fault model with time (10 s) (*62*). Inset shows the moment rate function of the finite fault model (gray shaded area) and the total source time function of the MPS solution, with focal mechanism solutions projected to the western hemisphere.

south of that in the initial United States Geological Survey (USGS) report (22.013°N) (https://www.usgs.gov/programs/earthquake-hazards/earthquakes), but in closer agreement with the Thailand catalog (21.68°N). The refined origin time (2025-03-28 06:20:55.209 coordinated universal time) is 2.5 s later than the USGS report, again aligning more closely with the Thailand catalog (https://www.tmd.go.th/en/EarthQuake).

The supershear rupture pulse recorded at the NPW station, identified as the onset of the largest amplitude in the velocity waveform, arrived at ~48 ± 0.5 s (Fig. 4C inset). Given the station's distance of 206 km from the refined epicenter—closely approximating the rupture distance along the fault—the inferred average rupture speed was ~4.3 km/s, exceeding the local shear-wave velocity and thus indicating supershear propagation. Notably, the rupture phase arrival time at NPW closely matches the rupture timing derived from BP and MPS results at the same location (blue triangle in Fig. 4C). Further evidence for supershear rupture propagation is provided by the NPW waveform characteristics. Theoretical asymptotic analyses and laboratory rock experiments (*50*) demonstrate that when the Mach cone front generated by a sustained supershear rupture sweeps a near-fault station, the resulting ground motion is dominated by fault-parallel (FP) particle velocity relative to the fault-normal (FN) component. This pattern is clearly observed at the NPW station (Fig. 4C, inset, and fig. S20A). Moreover, the observed FP-to-FN velocity amplitude ratio suggests a rupture velocity of 5.0 to 5.3 km/s (fig. S20B and methods), consistent with estimates from BP and MPS.

Using the refined epicenter and rupture speed constraints from the preceding analyses, we performed a finite fault inversion that jointly fits teleseismic body waves, near-fault strong-motion records, and the 3D surface displacement field (methods). The resulting finite fault model (Fig. 4D) provides excellent fits to all datasets (figs. S21 to S23) and yields a moment-rate function consistent with that derived from the MPS source time function (Fig. 4D, inset), lending high confidence to key features of the model. These include the fault-geometry transition from east-dipping in the north to west-dipping in the south, the spatial distribution of slip, and the supershear rupture velocity observed along the central segment of the rupture. Collectively, multiple independent analyses and inversions corroborate a sustained supershear rupture propagating southward from ~100 km to at least 310 km along the strike from the epicenter (Fig. 4).

Thick LVFZ for supershear segment

We used passive seismic measurements to investigate the influence of the internal structure of the Sagaing fault zone. Along the supershear rupture segment of the 2025 event, three broadband seismic stations were located within 1 km of the surface rupture (Fig. 5A and fig. S24). Teleseismic P-waves recorded at these stations exhibited major time shifts between the vertical and radial components (fig. S24A). Notably, earthquakes originating from the western side of the fault show larger time shifts than those from the eastern side, indicating a clear back-azimuthal variation (fig. S24B). These time shifts resulted in delayed arrivals in receiver function (RF) waveforms, and when the shift exceeded a threshold (e.g., >1s), a negative phase was observed at the zero time (Fig. 5B, figs. S25 to S27, and methods).

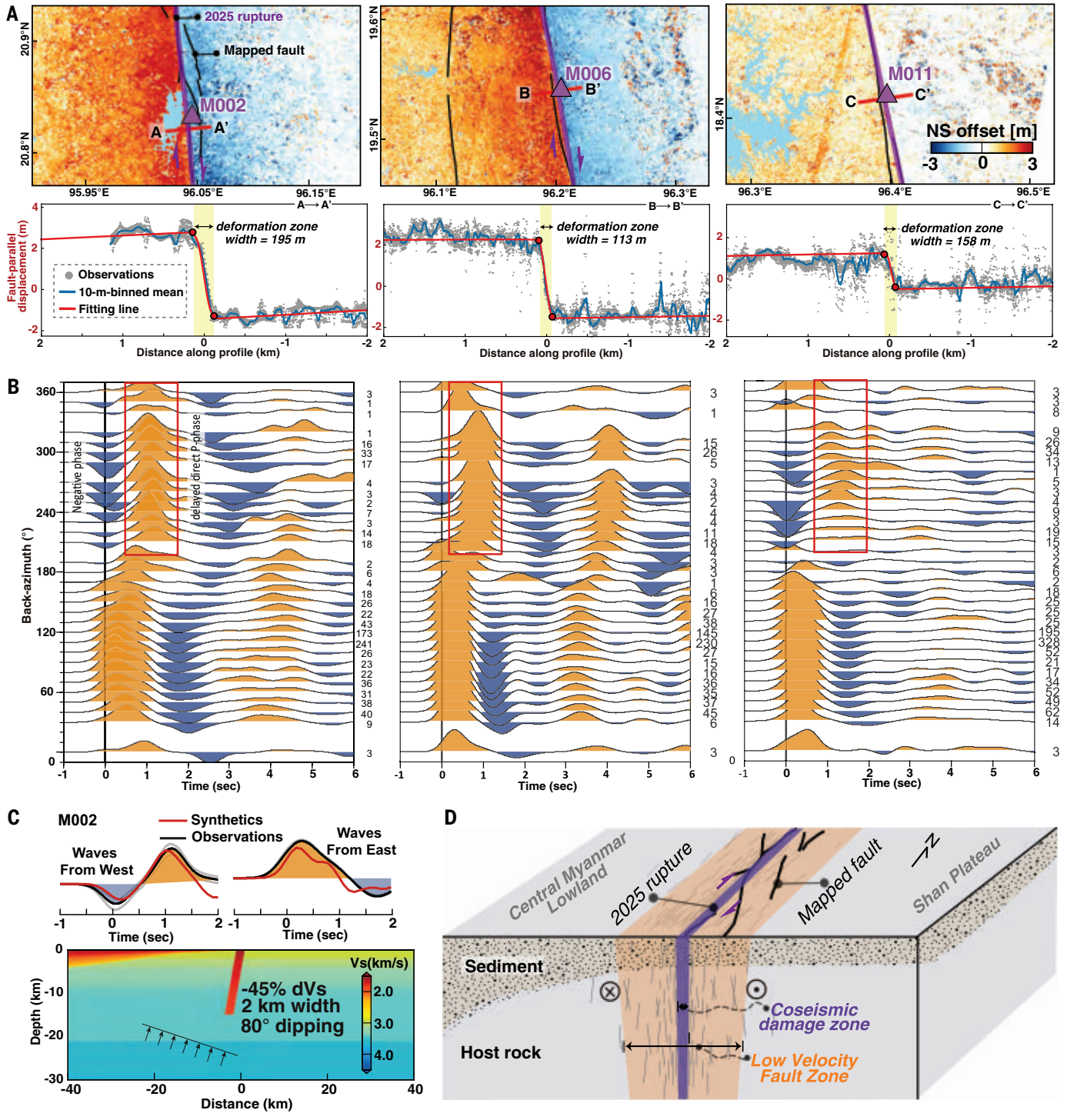


Fig. 5. Relationship between the coseismic damage zone and the low-velocity fault zone sustaining supershear rupture. (**A**) North-South displacement and across-fault profiles at the three broadband stations (M002, M006, and M011; Fig. 1C). Black and purple lines represent mapped faults and the 2025 rupture. (**B**) Stacked RF waveforms at M002, M006, and M011, grouped by the back-azimuth of incoming waves, with the number of stacked traces indicated on the right side of each panel. A clear transition is observed from a positive to negative zero-lag phase at back-azimuths near 180°. (**C**) 2D synthetic modeling results illustrating the effect of a dipping low-velocity fault zone on RF waveforms. A west-dipping fault zone produces a negative zero-lag phase for waves coming from the west, while maintaining a positive first P-phase for waves from the east, consistent with the observations. (**D**) Schematic illustration showing the relationship between the coseismic damage zone and the low-velocity fault zone.

These negative zero-lag phases in RFs cannot be explained by a 1D layered velocity model and are often neglected in conventional RF studies (*51*). Here, we argue that the negative zero-lag phase represents true structural signals caused by 2D effects of low-velocity fault zone structure. Through synthetic tests we show that basin models (*37*) are unable to reproduce the negative phase at time zero (fig. S28B). The negative zero-lag phase occurs only within specific back-azimuth ranges (Fig. 5B), which rules out sensor misorientation. It is observed when a different frequency band (fig. S27A) and deconvolution method (fig. S27B) are used. Therefore, the signal is not caused by sidelobe effects during filtering nor is it an artifact of RF calculation. The transition from a positive to a negative zero-lag phase occurred at back-azimuths near 180°—consistent with the polarity reversal observed in Fig. 5B—and closely aligns with the strike of the Sagaing fault, supporting an interpretation of a west-dipping fault zone structure.

We modeled the observed polarity reversal of the zero-lag phase with a west-dipping low-velocity fault zone structure (Fig. 5C and methods). A grid search was performed on the dipping angle, fault zone width, and velocity reduction, to find the best combined fit to observed RF waveforms. The velocity reduction ranged from 25 to 45% with 10% intervals; fault zone width varied from 1 to 3 km with 0.5 km intervals; and the dip angle ranged from 60° to 90° with 10° intervals. For each grid point, we calculated the amplitude ratio between the negative phase near time zero and the subsequent positive phase, then computed the misfit between observed and synthetic waveforms. The search shows that a west-dipping fault zone with a dip angle of 70° to 80°, a thickness of ~2 km, and a shear-wave velocity reduction of approximately 45% is needed to achieve the negative zero-lag phase and corresponding amplitude (figs. S28 and S29). The grid search results also indicate an uncertainty of ~0.5 km in fault width and ~10% in velocity reduction (fig. S30 and methods). The inferred west-dipping fault geometry is also consistent with that derived from geodetic inversions (fig. S31). The ~2-km width of the fault zone exceeds those of other well-studied faults such as the San Andreas, Ridgecrest, Landers, and Calico faults (table S3), although the observed velocity reduction is comparable (fig. S32 and methods).

The drivers of supershear rupture

The simple fault geometry and thick fault zone probably provide the most essential conditions behind the supershear and subshear-to-supershear transition distance of the 2025 event. A fault zone produces strong head waves that travel faster than the initial rupture front, creating dynamic fault stress perturbations that promote the development of a new supershear rupture front, as confirmed by dynamic earthquake simulations (*12*, *52*, *53*). Fault zones need to be sufficiently thick to facilitate a supershear transition. Damic rupture simulations show that the critical fault zone width should be at least 2 to 3 times the critical nucleation size of earthquakes $L_c = \frac{GD_c}{\Delta\tau_s}$, where G is shear modulus, D_c is the critical slip distance in slip-weakening friction law, and $\Delta\tau_s$ is the fault strength drop, i.e., the difference between fault static and dynamic shear strength. Assuming $G = 30$ GPa, $D_c = 1$ m, and $\Delta\tau_s = 50$ MPa (methods), and yielding a L_c of 600 m, the observed Sagaing fault zone width would be ~3 to 4 L_c. Such thick fault zones allow the transition to supershear rupture even with low background stress. Given the estimated static stress drop of 3 MPa, the relative stress drop—defined by the ratio of static stress drop to the strength drop—is likely <0.1 in this region, substantially lower than the relative stress drop needed for supershear transition in homogeneous media (*12*). Another feature of supershear rupture in thick fault zones is the increase in the supershear transition length with the fault zone width (*12*). A relative stress drop of <0.1 and a fault zone width of 3 to 4 L_c indicate that the transition length is >100 L_c (*12*), which matches the ~100-km transition length observed from the seismological analysis.

Other factors may also have facilitated the supershear rupture transition in the 2025 event. The observed supershear transition occurred near an along-strike fault-geometry transition, where the fault dip changes from eastward to subvertical (slightly west-dipping, Fig. 3B and fig. S5). Dynamic simulations demonstrated that geometrical kinks, such as dip changes or bends, can act as stress concentrators that temporarily slow and then reaccelerate the rupture, favoring supershear transition once the accumulated energy is released (*54*). As the rupture propagated onto the west dipping fault, the smooth fault geometry provided the "superhighway" to help sustain the supershear rupture, similar to what may have happened for other supershear events (*9*, *11*). The relatively thick (~2 to 3 km) sedimentary basin along the southern fault segment (Fig. 3E) serves as another favorable condition to sustain supershear rupture, as the waves reflected from the bottom of the basin could have increased shear stress on the fault and promoted rupture propagation (*55*, *56*).

In contrast to the relatively simple surface rupture of the 2025 event, fault traces mapped by earthquake geologists reveal secondary fault branches and bifurcations (Figs. 5D and 1C). These deviations from the straight 2025 rupture likely reflect past earthquakes that created more complex geomorphological signatures (*27*). The considerable fault zone thickness (~2 km) further suggests that ruptures in the past may have involved both the main fault and secondary strands. Observations indicate that fault zones may take hundreds or thousands of years for faults to heal from coseismic damages of large events (*57*). Dynamic simulations also show that coseismic rupture can produce off-fault plastic deformation that leads to permanent damage in the fault zone (*58*). Therefore, cumulative fault damage from previous events, including events similar to the 2025 event, have likely formed the observed thick fault zone (Fig. 5D).

Our observations and models reconcile the coseismic fault geometry, supershear rupture, and fault damage zone in a region where supershear rupture may have been governed by the cumulative imprint of past earthquakes. This inherited complexity appears to have guided the rupture dynamics for the 2025 M_w 7.8 Mandalay event. The Sagaing fault thus exemplifies how long-term fault zone evolution may influence earthquake behavior, a pattern that might apply to other plate boundary–type strike-slip faults. Specifically, supershear rupture along these faults may be more likely to occur where the fault zones are wide and fault geometry is simple. These findings highlight the need for both specific-event and earthquake-cycle models to incorporate more comprehensive and accurate fault zone structures (*59*), which may critically control rupture mode, speed, and size across successive earthquake generations. The results also provide new insight into assessing the potential hazard of major earthquakes along continental faults with recognized seismic gaps, such as those associated with the southern San Andreas fault in CA and the North Anatolian fault near İstanbul, Türkiye.

REFERENCES AND NOTES

1. M. Bouchon, M. Vallée, *Science* **301**, 824–826 (2003).
2. Y. Klinger *et al.*, *Bull. Seismol. Soc. Am.* **95**, 1970–1987 (2005).
3. K. T. Walker, P. M. Shearer, *J. Geophys. Res. Solid Earth* **114**, 2008JB005738 (2009).
4. E. M. Dunham, R. J. Archuleta, *Bull. Seismol. Soc. Am.* **94**, S256–S268 (2004).
5. H. Zeng, S. Wei, A. Rosakis, *J. Geophys. Res. Solid Earth* **127**, e2022JB024359 (2022).
6. D. Melgar *et al.*, *Seismica* **2**, v2i3.387 (2023).
7. C. Ren *et al.*, *Science* **383**, 305–311 (2024).
8. Z. Jia *et al.*, *Science* **381**, 985–990 (2023).
9. M. Bouchon *et al.*, *Tectonophysics* **493**, 244–253 (2010).
10. J. Jara *et al.*, *Proc R Soc Lond A.* **477**, 20210364 (2021).
11. D. P. Robinson, S. Das, M. P. Searle, *Tectonophysics* **493**, 236–243 (2010).
12. Y. Huang, J.-P. Ampuero, D. V. Helmberger, *Earth Planet. Sci. Lett.* **433**, 109–115 (2016).
13. A. Vallage, Y. Klinger, R. Grandin, H. S. Bhat, M. Pierrot-Deseilligny, *Geology* **43**, 1079–1082 (2015).
14. K. Okubo *et al.*, *J. Geophys. Res. Solid Earth* **124**, 11771–11801 (2019).

15. C. W. D. Milliner *et al.*, *Geochem. Geophys. Geosyst.* **16**, 1577–1598 (2015).
16. S. L. Antoine, Y. Klinger, K. Wang, R. Bürgmann, *Geophys. Res. Lett.* **51**, e2024GL110798 (2024).
17. X. Xi *et al.*, *Commun. Earth Environ.* **6**, 103 (2025).
18. C. Li, T. Li, X. Shan, G. Zhang, *Seismol. Res. Lett.* **94**, 39–51 (2022).
19. Y. Ben-Zion, C. G. Sammis, *Pure Appl. Geophys.* **160**, 677–715 (2003).
20. C. Pelties, J. de la Puente, J.-P. Ampuero, G. B. Brietzke, M. Käser, *J. Geophys. Res. Solid Earth* **117**, 2011JB008857 (2012).
21. C. Li, Z. Ma, X. Xi, G. Zhang, X. Shan, *Geophys. Res. Lett.* **51**, e2024GL108169 (2024).
22. Y. Huang, J.-P. Ampuero, *J. Geophys. Res. Solid Earth* **116**, 2011JB008684 (2011).
23. Y. Huang, J.-P. Ampuero, D. V. Helmberger, *J. Geophys. Res. Solid Earth* **119**, 3133–3154 (2014).
24. M. Abdelmeguid, M. S. Mia, A. Elbanna, *Geophys. Res. Lett.* **51**, e2023GL108060 (2024).
25. V. Lyakhovsky, Y. Ben-Zion, A. Agnon, *J. Geophys. Res.* **106**, 4103–4120 (2001).
26. L. Valoroso, L. Chiaraluce, C. Collettini, *Geology* **42**, 343–346 (2014).
27. Y. Wang, K. Sieh, S. T. Tun, K.-Y. Lai, T. Myint, *J. Geophys. Res. Solid Earth* **119**, 3767–3822 (2014).
28. E. O. Lindsey *et al.*, *Earth Planet. Sci. Lett.* **622**, 118384 (2023).
29. N. Hurukawa, P. Maung Maung, *Geophys. Res. Lett.* **38**, 2010GL046099 (2011).
30. Earthquake survivors in Mandalay face sanitation crisis in temporary shelters (Mizzima, 2025); https://bur.mizzima.com/2025/04/09/52626.
31. Strong earthquake in central Myanmar felt as far as Bangkok (Reuters, 2025); https://www.reuters.com/world/asia-pacific/strong-earthquake-central-myanmar-panic-bangkok-2025-03-28.
32. Myanmar earthquake caused damage to 847 households in Ruili, Yunnan [in Chinese] (Xinhua News Agency, 2025); https://www.xinhuanet.com/20250329/b3bcf429bdb544bb966b28c93a96bd5a/c.html.
33. N. Inoue *et al.*, *Seismica* **4**, v4i1.1691 (2025).
34. D. Melgar *et al.*, *Seismica* **4**, v4i2.1771 (2025).
35. Z. Peng *et al.*, Earthquake Research Advances, 100413 (2025).
36. S. L. Antoine, Y. Klinger, A. Delorme, R. D. Gold, *J. Geophys. Res. Solid Earth* **127**, e2022JB024480 (2022).
37. Y. Bai *et al.*, *Earth Planet. Sci. Lett.* **610**, 118105 (2023).
38. R. S. Stein, P. Bird, *Seismol. Res. Lett.* **95**, 3406–3415 (2024).
39. Y. Klinger, *J. Geophys. Res. Solid Earth* **115**, 2009JB006550 (2010).
40. H. S. Bhat, M. Olives, R. Dmowska, J. R. Rice, *J. Geophys. Res. Solid Earth* **112**, 2007JB005027 (2007).
41. P. Bird, R. S. Stein, *Seismol. Res. Lett.* **95**, 3306–3315 (2024).
42. G. Bertrand *et al.*, *Comptes Rendus Académie Sci. - Ser. IIA - Earth. Planet. Sci.* **327**, 479–484 (1998).
43. Y. Fialko, D. Sandwell, M. Simons, P. Rosen, *Nature* **435**, 295–299 (2005).
44. X. Xu *et al.*, *Geophys. J. Int.* **204**, 1867–1886 (2016).
45. Z. Jin, Y. Fialko, *Geophys. Res. Lett.* **48**, e2021GL095213 (2021).
46. Z. Ma *et al.*, *Geophys. Res. Lett.* **51**, e2023GL107788 (2024).
47. K. Jiang, W. Xu, L. Xie, *J. Geophys. Res. Solid Earth* **129**, e2024JB028826 (2024).
48. Q. Shi, S. Wei, M. Chen, *Geophys. J. Int.* **215**, 737–752 (2018).
49. S.-T. Lai *et al.*, *Earth Syst Sci Data Discuss* **2025**, 1–23 (2025).
50. M. Mello, H. S. Bhat, A. J. Rosakis, *J. Mech. Phys. Solids* **93**, 153–181 (2016).
51. X. Yang, G. L. Pavlis, Y. Wang, *Bull. Seismol. Soc. Am.* **106**, 1948–1962 (2016).
52. S. P. Hicks *et al.*, *Nat. Geosci.* **13**, 647–653 (2020).
53. H. Weng, J.-P. Ampuero, *J. Geophys. Res. Solid Earth* **124**, 8584–8610 (2019).
54. L. Bruhat, Z. Fang, E. M. Dunham, *J. Geophys. Res. Solid Earth* **121**, 210–224 (2016).
55. H. Weng, H. Yang, Z. Zhang, X. Chen, *J. Geophys. Res. Solid Earth* **121**, 4446–4457 (2016).
56. Y. Kaneko, N. Lapusta, *Tectonophysics* **493**, 272–284 (2010).
57. E. S. Cochran *et al.*, *Geology* **37**, 315–318 (2009).
58. P. Zhai, Y. Huang, C. Liang, J.-P. Ampuero, *Geophys. J. Int.* **242**, ggaf274 (2025).
59. Y. Ben-Zion, *Rev. Geophys.* **46**, 2008RG000260 (2008).
60. W. Fadil *et al.*, *Bull. Seismol. Soc. Am.* **113**, 613–635 (2023).
61. L. Y. Zhang *et al.*, *Geochem. Geophys. Geosystems* **21**, e2020GC009091 (2020).
62. S. Wei, Data and model files of "Supershear Rupture Sustained Through a Thick Fault Zone in the 2025 Mw 7.8 Mandalay Earthquake", Zenodo (2025); https://doi.org/10.5281/zenodo.17077559.

ACKNOWLEDGMENTS

We thank Y. Wang of the Department of Geology, National Taiwan University for the discussion on the earthquake geology and acknowledge Center for Geohazard Observations and Department of Meteorology and Hydrology of Myanmar for their work in seismic data acquisition. We thank the China Earthquake Networks Center for providing continuous waveform data. We acknowledge the Department of Meteorology and Hydrology in Naypyitaw and the GFZ Helmholtz Center for Geosciences (GFZ) for their collaborative efforts to install the strong motion station at Naypyitaw and make the data available. **Funding:** Funding for this work was provided by the Chinese Academy of Sciences (grant E4515401), the National Natural Science Foundation of China (grants 42494861 and NSFC92355001), the Postdoctoral Fellowship Program of CPSF (grant GZB20250098), the Tier 2 project "Integrating space geodesy and dense array seismology to reveal fundamental earthquake physics" supported by the Singapore Ministry of Education (MOE) (MOET2EP50124-0011). **Author contributions:** Conceptualization: S.W., X.W., Z.M.; Methodology: S.W., X.W., Z.M., C.L., H.Z., Q.S., H.C., M.L., Y.H.; Investigation: S.W., Z.M, X.W., C.L., H.Z., Q.S., H.C., M.L., Y.H.; Visualization: Z.M., C.L., X.W., H.C., S.W., Q.S., H.Z., Y.B., S.Y.; Funding acquisition: S.W., X.S.; Project administration: S.W.; Supervision: S.W.; Writing – original draft: S.W., M.Z., X.W., H.Z., C.L., Q.S.; Writing – review & editing: All authors. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** The ALOS-2 PALSAR images are owned by the Japanese Space Agency (JAXA) and provided to Z.M. under grant ER4A2N150. All Sentinel-1 SAR images are provided by the European Space Agency (ESA). All Sentinel-2 optical images are downloaded from the Google Earth engine applications (https://code.earthengine.google.com/_ah/conflogin?continue=https://code.earthengine.google.com/). The teleseismic data and strong motion waveforms are from the IRIS-DMC. Our SAR/InSAR. Part of earthquake source parameters (e.g., mainshock epicenter, moment tensor solutions) are downloaded from https://www.globalcmt.org/CMTsearch.html and https://www.usgs.gov/programs/earthquake-hazards. 3D surface deformation map, multiple-point-source solution, back-projection results, finite fault model, static slip model, total slip along the rupture and damage zone width can be downloaded from Zenodo (*62*). Strong motion waveform data were downloaded from strongmotioncenter.org. The raw teleseismic waveform data from the Earth Observatory of Singapore–Myanmar (EOS–Myanmar) seismic array and the processed receiver function (RF) waveforms are available for download at the World Data Center for Geophysics, Beijing (http://www.geophys.ac.cn/ArticleDataInfo.asp?MetaId=727). **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/content/page/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adz2101
Methods and materials; figs. S1 to S32; Tables S1 and S3; References (*63*–*122*)

Submitted 22 May 2025; accepted 8 September 2025

10.1126/science.adz2101

RESEARCH ARTICLE

Bimaterial effect and favorable energy ratio enabled supershear rupture in the 2025 Mandalay earthquake

Liuwei Xu[1], Lingsen Meng[1]*, Zhang Yunjun[2,3], Yanchen Yang[4,5], Yidi Wang[2,3], Changyang Hu[2,3], Huihui Weng[6], Wenbin Xu[7], Elizabeth Su[1], Chen Ji[8]

Joint seismic and geodetic analyses revealed that the 2025 moment magnitude (M_w) 7.8 Mandalay, Myanmar, earthquake ruptured ~510 km of the Sagaing fault, with a sustained supershear rupture extending ~450 km on the southern branch. Far-field Mach waves and near-field ground motion confirmed the supershear nature. This exceptionally long supershear rupture caused building collapse and soil liquefaction, as observed in satellite imagery, offering insights into the damage potential of such ruptures in urban areas. Sustained supershear propagation was facilitated by the fault's linear geometry, prolonged interseismic quiescence, favorable energy ratio, and pronounced bimaterial contrasts across the fault interface. These findings underscore the roles of fault structure, stress accumulation, and material contrasts in governing rupture dynamics, demonstrating that large-scale supershear propagation can occur in interplate continental fault systems.

At 12:51 p.m. local time on 28 March 2025, a moment magnitude (M_w) 7.8 earthquake struck near the city of Mandalay, Myanmar. Situated at the junction of three tectonic plates—the Indian (IN), Eurasian (EU), and Burma (BU) plates (*1, 2*)—Myanmar is subject to high seismicity. The central longitudinal axis of the country, home to both the major economic center of Mandalay and the political capital NayPyiDaw, coincides with the boundary between the BU and EU plates (Fig. 1A). The right-lateral strike-slip Sagaing fault, which marks the boundary, spans ~1400 km and accommodates a relatively high slip rate of ~20 mm/year (*3*, *4*). In the 20th century, the Sagaing fault generated at least seven magnitude (M) ≥ 7 earthquakes along different segments (Fig. 1A), with the exception of a seismic gap between Mandalay and NayPyiDaw (Fig. 1A) (*2*, *5*). The 2025 Mandalay earthquake ruptured this long-standing seismic gap. The epicenter was located near Mandalay, and early reports indicated a bilateral, north-south (N-S)–oriented rupture (*6*). According to data compiled by the Democratic Voice of Burma, the earthquake resulted in 4355 confirmed fatalities, 210 people reported missing, and 7830 injured.

To obtain timely and comprehensive fault rupture details about the earthquake, we collected global seismic data and applied the slowness-enhanced back-projection (SEBP) method (*7*). SEBP enabled us to resolve the spatiotemporal evolution of high-frequency (HF) radiators along the fault during the mainshock. We also derived coseismic ground deformation using synthetic aperture radar (SAR) and optical imagery from the Sentinel-1, LuTan-1, ALOS-2, and Sentinel-2 satellites. These deformation observations, in combination with global and local seismic datasets, were jointly inverted through finite fault inversion (FFI) (*8*, *9*) to reconstruct the slip distribution, rupture history, and slip vectors across the fault. Our integrated analysis of SEBP and joint FFI results revealed that the earthquake ruptured the subvertical Sagaing fault in a bilateral manner along the N-S direction. The northern branch rupture was relatively minor, with an approximate length of 60 km and an average rupture speed of 0.9 km/s. By contrast, the southern branch exhibited an unusually fast rupture, propagating ~450 km at a speed of 5.0 km/s, exceeding the local shear wave velocity and therefore classifying it as a supershear rupture. We confirmed the supershear nature of the southern segment by identifying Mach waves and Mach cones in the seismic wavefield and by analyzing ground motions recorded at a near-fault station.

Owing to the ongoing civil conflict in Myanmar, field investigations and postevent damage assessments have been substantially constrained. To address this limitation, we generated damage proxy maps (DPMs) and analyzed them in conjunction with other satellite radar imagery. DPMs highlight regions where the radar backscattering characteristics changed anomalously relative to their typical background variation, indicative of potential structural damage (*10*, *11*). Optical satellite images revealed building collapses in urban areas and widespread secondary hazards (e.g., soil liquefaction, landslides), which aligned with regions of high DPM values. This study offers a detailed characterization of the prolonged and devastating earthquake, providing a set of observations that can serve as a reference for postevent hazard assessment.

Rupture kinematics resolved by SEBP and joint FFI

The SEBP analysis revealed an asymmetrical bilateral rupture propagating in the N-S direction. The HF radiators closely followed the surface trace of the Sagaing fault (Fig. 1A and figs. S1 to S9). All three arrays consistently indicated a fast rupture speed of 4.8 to 5.0 km/s on the southern branch (Fig. 1C; figs. S2B, S4B, and S6B; and movies S1 to S3). The length of the southern rupture segment was ~450 km. The speed substantially exceeded the local shear wave speed in the crust above depths of 20 km (2.5 to 3.7 km/s; fig. S10 and table S1) (*5*), suggesting supershear rupture. By contrast, the northern branch exhibited a shorter and slower rupture. Based on SEBP results, the northern rupture extended ~60 km with an estimated average speed of 0.9 km/s (*12*). The rupture speeds resolved by the three arrays showed greater variability, ranging from 0.9 to 1.7 km/s (figs. S2B, S4B, and S6B). This variation likely reflects uncertainties introduced by the relatively weak HF radiation from the northern branch, attributable to its slower rupture speed and interference from the southern branch's signal (*13*). We validated the speed measurements by applying back-projection to synthetic seismograms generated by a bilateral model with a subshear northern branch and a supershear southern branch (*12*) (figs. S11 to S13). The recovered southern speeds were 8 to 20% lower than the input value, whereas those of the northern branch were underestimated by 30% (figs. S11 and S12). Although such underestimation may be less pronounced in real earthquake back-projection, it suggests greater uncertainty in back-projection results for the northern branch. Nevertheless, the synthetic back-projection confirmed the supershear nature of the southern rupture and the subshear nature of the northern one.

The joint FFI results indicated that the rupture was predominantly right-lateral strike-slip (Fig. 2C and movie S4). The coseismic slip distribution aligned well with the HF radiators (Fig. 1C); north of the epicenter, slip extended ~60 km, whereas the southern rupture extended

[1]Department of Earth, Planetary and Space Sciences, University of California Los Angeles, Los Angeles, CA, USA. [2]National Key Laboratory of Microwave Imaging, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. [3]School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing, China. [4]National Engineering Research Center for Remote Sensing Satellite Applications, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. [5]School of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China. [6]State Key Laboratory for Mineral Deposits Research, School of Earth Sciences and Engineering, Nanjing University, Nanjing, China. [7]School of Geoscience and Info-Physics, Central South University, Changsha, China. [8]Department of Earth Science, University of California, Santa Barbara, Santa Barbara, CA, USA. ***Corresponding author. Email: lsmeng@g.ucla.edu**

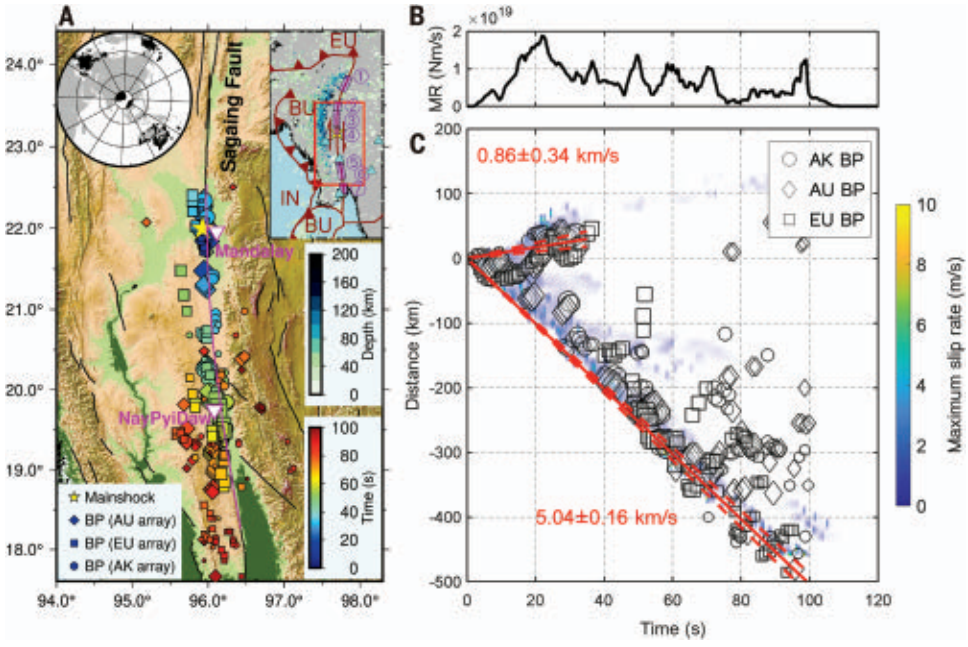


Fig. 1. Overview map and rupture process. (**A**) Back-projection (BP) results and active faults (black lines). (*2*). Colored symbols represent HF radiators imaged by three seismic arrays (see legend). Symbol size is proportional to beam power, and color indicates rupture time relative to the mainshock origin time (the lower legend). White triangles denote the major cities along the Sagaing fault. The magenta line denotes the surface rupture identified according to the ground deformation. (Inset, upper left) Locations of the three arrays. (Inset, upper right) The regional tectonic map. Colored dots indicate historical seismicity from 1 January 1990 to 3 April 2025, color-coded by hypocenter depth (the upper legend). Dark red lines mark plate boundaries (*1*, *2*). The red box highlights the source region shown in the main image. Cyan triangles denote strong motion stations. Purple bars denote rupture zones of historical *M* 7 earthquakes in the 20th century, with event names and magnitudes listed in table S3 (*2*). BU, Burma plate; IN, Indian plate; EU, Eurasian plate. (**B**) Moment rate function illustrating the temporal evolution of seismic moment release. (**C**) Rupture velocity from SEBP and comparison to the FFI. Symbols mark the timing and location of HF radiators. Locations are shown as along-fault distance relative to the hypocenter, with positive values to the north and negative values to the south. The red solid line indicates the best-fit rupture front, with standard deviation (±1 σ) outlined by dashed lines. The background colormap represents the maximum slip rate at each along-strike position from the FFI.

~450 km. Near the hypocenter, the fault dips 65° to the east, whereas along the southern branch, it steepens to 80° [see also (*12*)]. Slip was concentrated at shallow depths: About 75% of the seismic moment was released between 0- and 15-km depth, with a peak slip of ~7.9 m. The total seismic moment was 6.37×10^{20} N·m, corresponding to a moment magnitude of M_w 7.8. Most of the moment release occurred within the first 100 s (Fig. 1B). Rupture duration exceeded the typical value predicted by the global scaling relationship [~40 s (*14*)] but was comparable to that of some M_w 7.8 strike-slip events, such as the 2001 Kunlun (*15*), 2013 Scotia (*16*), and 2023 Türkiye earthquakes (*17*). Our model successfully reproduced static displacements observed in SAR and optical satellite data (figs. S14 to S16) and provided good fits to teleseismic and local strong-motion recordings (figs. S17 to S22), including at station NPW, located just 2 to 3 km from the fault (figs. S21 and S22). Nine additional inversions using different initial models yielded minor standard deviations in both objective function values and coseismic slip distributions [(*12*); fig. S23], confirming the robustness of our solution.

In the back-projection images, we observed that the HF radiators extended farther along the southern branch than mapped surface fault traces (Fig. 1A). The slip model revealed that the southern segment between 420 and 460 km hosted coseismic slip that did not reach the surface, thereby accounting for the discrepancy in rupture extent. Additionally, several southern radiators observed after 90 s exhibited a westward spatial bias relative to the fault trace, likely caused by coda-wave contamination during the rupture's terminal phase. The distribution of northern HF radiators was consistent with both the mapped surface traces and the slip model after calibration (*12*). The northern surface rupture terminated near ~22.5°N, coinciding with the southern edge of a M_w 6.8 strike-slip earthquake in 2012 (fig. S24) (*18*). This termination may reflect prior stress release from the 2012 event that impeded rupture propagation.

Aftermath estimation by satellite damage proxy maps

The DPMs derived from Sentinel-1A satellite imagery have served as a key tool for systematically assessing the mainshock and its impacts in the absence of on-ground postevent damage assessments (Fig. 3) (*10*–*12*, *19*). We found that areas exhibiting high DPM values closely align with the spatial orientation of fault traces, consistently forming a 2- to 3-km-wide zone along the fault (Fig. 3 and fig. S25). This pattern helps delineate the extent of off-fault damage and reveals widespread surface destruction beyond the primary rupture zone. By integrating building footprint data with hazard models from the US Geological Survey (USGS) and using a causal Bayesian network to model seismic multihazards and their impacts (*12*, *20*), we generated three high-resolution (45 m) hazard probability maps for building damage, liquefaction, and landslides, respectively (Fig. 3 and figs. S26-S28). Building collapses were observed in Mandalay, with high landslide potential identified at Sagaing Hill and high soil liquefaction potential in the Irrawaddy floodplain and the city's southeastern areas (Fig. 3). These findings underscore the multiscale mechanisms underlying seismic impacts and demonstrate the effectiveness of satellite-based seismic multihazard and impact assessments, particularly in regions where on-the-ground access is limited.

Validation of supershear rupture

We further validated the southern supershear rupture using surface wave observations. As proposed by (*21*), a supershear rupture exhibits distinct characteristics in the far-field surface wave field: within specific regions, waves from different parts of the rupture arrive simultaneously and interfere constructively, simplifying the Rayleigh and Love waveforms. This results in a high degree of similarity between the waveforms of the large supershear rupture and those of a smaller, colocated event (hereinafter referred to as an empirical Green's function, EGF). These regions of constructive interference are known as Mach cones, with the recorded surface waves referred to as Mach waves, and the angles between the Mach cones and the rupture direction termed Mach angles (*12*). We collected 76 Rayleigh waveforms to identify Mach waves and cones. To minimize surface wave dispersion, seismograms were bandpass filtered to 15 to 25 s following (*21*, *22*). The EGF was a M_w 5.1 event that occurred north of the mainshock epicenter with a similar focal mechanism (EGF 2, table S2 and fig. S24). We quantified waveform similarity using standard cross-correlation

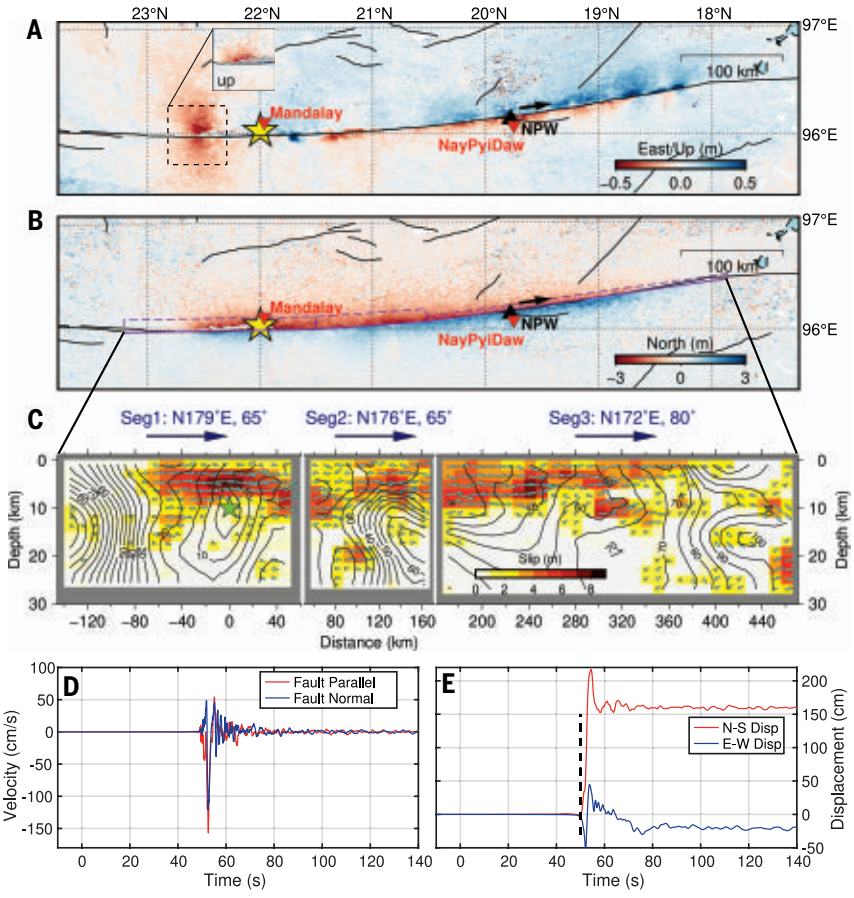


Fig. 2. Ground deformation and slip model. (**A**) East-west (E-W) ground deformation. The zoomed-in inset presents the vertical ground deformation with full coverage shown in fig. S32C. Inverted triangles denote major cities along the fault. The black triangle indicates the seismic station GE.NPW, and the arrow shows the FP direction (N172°E). The deformation reference point is set at (N20.0°, E94.6°), outside the figure extent. (**B**) N-S ground deformation. Purple boxes outline the boundaries of the fault model used in the joint FFI, with solid lines denoting the upper fault boundaries and dashed lines representing other boundaries. (**C**) Spatial distribution of final slip (color scale), rupture initiation time (black contours in seconds), and rake angle (cyan arrows) along the fault. Strike and dip angles for each fault segment are labeled at the top. (**D**) Comparison of FP (red) and FN (blue) velocity seismograms recorded at GE.NPW. Waveforms are obtained by integrating the raw acceleration data and removing baseline drifts. The time axis is referenced to the earthquake origin time (28 March 2025, 06:20:52 UTC). (**E**) Ground displacement time series at GE.NPW. N-S (red) and E-W (blue) components are shown. The bold dashed line marks the onset of substantial ground deformation, occurring 50 s after the origin time. Displacements are calculated by integrating the velocity seismograms and applying drift corrections.

coefficients (CCs) (*12*). On the southeast side of the epicenter, waveforms from the mainshock and the EGF exhibited high similarity (CCs > 0.85) at several Australian stations (e.g., AU.GVL, AU.STKA; Fig. 4C). On the southwest side, Mach waves were observed at stations in South Africa and Madagascar (e.g., II.SUR, II.ABPO; Fig. 4C). The Mach angle was 30° to 60°, which, assuming a local S wave speed (V_s) of 3.2 km/s, suggests a rupture speed of 3.4 to 5.8 km/s. This range is consistent with the rupture speed measured by SEBP, though minor uncertainties may arise from V_s variations, rupture complexity, and rupture speed fluctuations.

Theoretical analyses and laboratory experiments (*23*, *24*) have shown that a key characteristic of supershear ruptures is a dominant fault-parallel (FP) component in the near-fault ground velocity field, which exceeds the fault-normal (FN) component. This FP dominance has been observed in several reported supershear events, including the 2002 Denali, 2020 Caribbean, and 2023 Türkiye earthquakes (*25*–*29*). The seismic station GE.NPW, located 2 to 3 km from the fault trace (Fig. 2, A and B), recorded ground accelerations during the mainshock. After integrating the acceleration to velocity and removing the drift (*30*), we observed a larger peak ground velocity in the FP direction (~160 cm/s) than in the FN direction (~120 cm/s) (Fig. 2D). This amplitude pattern is consistent with the aforementioned supershear signature. Recent observations on the 2023 Türkiye earthquake (*31*) showed that a rupture propagating at 3 km/s could also generate the FP dominance, with this speed exceeding the V_s of 2.83 km/s in their model. Dynamic simulations (*32*) further demonstrated that an FP/FN > 1 indicates supershear rupture once the front stabilizes after the transition, though this ratio may fluctuate if the rupture front remains unsteady immediately following the supershear onset. To provide further constraint on rupture speeds, we integrated the acceleration waveforms twice to obtain displacement time series (Fig. 2E). Near-fault displacement records are known to be sensitive to slip on nearby fault segments (*33*). We identified the onset of significant horizontal ground displacement and interpreted it as the approximate arrival of the rupture front at the fault patch near the station. Motions in the N-S component began as early as ~50 s (Fig. 2E and fig. S29), corresponding to an estimated rupture speed of 4.9 km/s, given the station's epicentral distance of 246 km. This represents a lower bound on the average rupture velocity from the hypocenter, as we ignored undetectable fault slip, the 2- to 3-km distance from the fault to the station, and the vertical rupture propagation within the fault. Taken together, these complementary observations, including SEBP imaging, far-field Mach wave detection, near-field amplitude characteristics, and displacement-based rupture timing, collectively support the existence of a fast-propagating supershear rupture along the southern segment of the fault.

Supershear ruptures controlled by bimaterial effect, energy ratio, and fault geometry

The 2025 Mandalay earthquake exhibited two remarkable features: an exceptionally early transition to supershear rupture and a prolonged supershear propagation spanning at least 400 km, making it one of the longest documented supershear earthquakes to date. Several mechanisms have been proposed to explain supershear transition, i.e., how rupture accelerates to the supershear regime. The classic Burridge-Andrews mechanism (*34*, *35*) suggests that the rupture front at a sub-Rayleigh speed is preceded by a daughter crack, whose speed is the S wave speed at the beginning and accelerates to a stable supershear speed after a certain distance. Given a high initial shear stress and/or a weak fault, rupture can also directly transition to the supershear regime before apparent formation of a daughter crack (*36*). In addition to the assumption of uniform properties, the heterogeneities of fault strength and stress (*36*–*38*), free surface (*39*), fault roughness, curvature, or bending (*40*, *41*), fault step-overs (*42*), damage zones (*43*), and bimaterial effects (*44*–*46*) may also promote supershear transition. The southern segment of the Mandalay earthquake shared several similarities with the 2018 M_w 7.5 Palu event but exhibited a higher rupture speed (5 km/s compared with 4.1 km/s in the Palu event). This exceptionally long and fast rupture offers a valuable opportunity to assess theoretical and numerical models of supershear propagation.

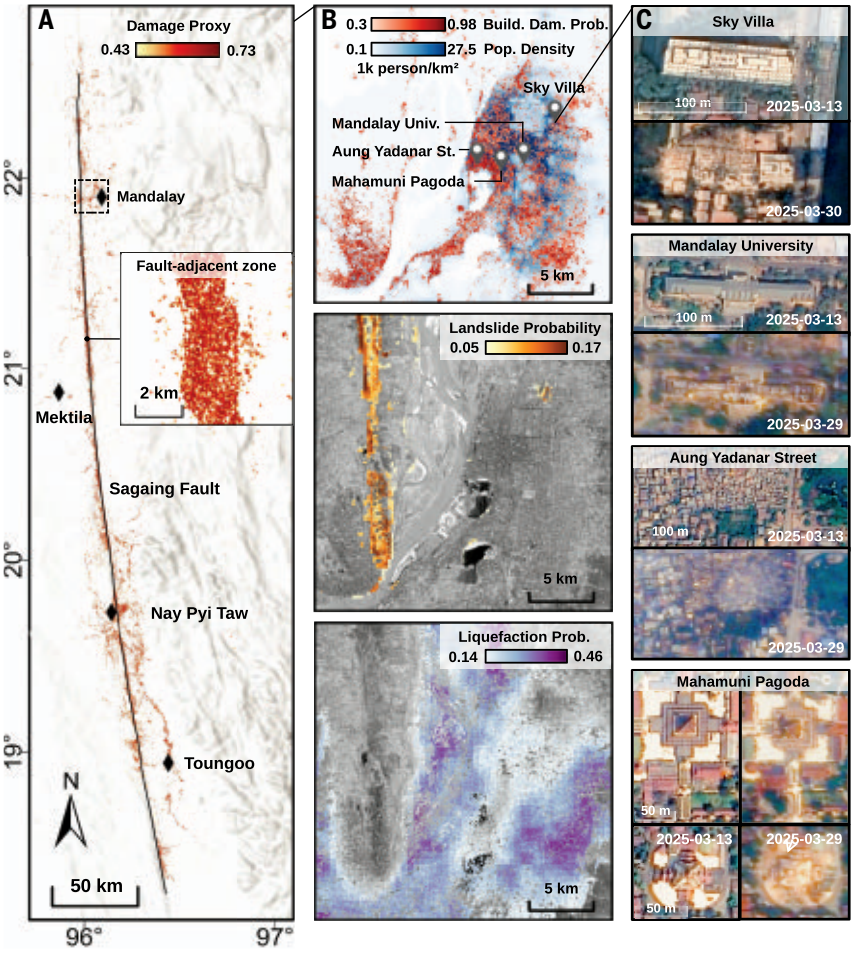


Fig. 3. Multihazard analysis of the earthquake region along the Sagaing fault. (**A**) DPM with white-to-red scale indicating increasing damage probability along the fault. (Inset) An enlarged view of a high-DPM zone near the fault. (**B**) Hazard probability distribution in Mandalay (light-to-dark scale), including building damage (top), landslides (middle), and liquefaction (bottom). Background layers show population density (white-to-blue scale) and Gaofen-1 optical imagery. (**C**) Damage interpretation in Mandalay derived from Jilin-1 optical imagery.

The Sagaing fault separates the Central Myanmar Basin to the west from Shan-Thai Block to the east (figs. S30 and S31). The former is characterized by a thick (~15 km) low-velocity layer consisting of Cenozoic basin sediments, and the latter features granite and the Mogok metamorphic belts with minimal sediment cover (*5*, *47*). The V_s structure along the Sagaing fault (*5*) reveals a pronounced material contrast across the fault within the shallowest 10 km, where most coseismic slip occurred. The eastern side exhibits V_s values of 2.8 to 3.7 km/s, whereas the western side shows lower V_s values of 2.2 to 3.3 km/s. By estimating the average velocity within the top 10 km as V_{s_mean}=10/Σ(dz/V_s), we obtained a V_{s_mean} of 2.9 and 3.3 km/s for the western and eastern sides, respectively. The difference in V_s suggests a difference in rock stiffness (*48*), a bimaterial contrast that influences rupture speed (*44*–*46*, *49*). Given the right-lateral strike-slip nature of the Sagaing fault, the southward rupture propagated in the negative direction, i.e., the moving direction of the stiff medium. Laboratory experiments and numerical simulations have shown that, in this direction, normal stress initially increases behind the rupture tip and decreases ahead of it. As the rupture propagates, the normal stress reduction becomes more pronounced, causing a broader region ahead of the rupture tip to yield simultaneously and accelerating the rupture to supershear speed. Once the rupture reaches supershear, the normal stress perturbation behind the crack tip reverses and becomes extensional rather than compressional (*44*–*46*, *50*). Conversely, rupture propagating in the positive direction (the moving direction of compliant medium) tends to experience increased normal stress ahead of the rupture tip, which favors subshear or nonsustained supershear behavior, explaining the slow rupture observed along the northern segment. However, dynamic simulations also predict a longer rupture extension in the positive direction compared with that in the negative direction (*51*, *52*), contrasting with the shorter northern branch of the Mandalay earthquake. This discrepancy could be explained by the five $M \geq 6.8$ events that occurred in the north between 1946 and 2012, which reduced stress accumulation (*2*) (Fig. 1A, table S3).

The observed rupture speed of 5 km/s was anomalously fast, corresponding to rupture speed/shear wave speed (V_r/V_s) ratios of 1.6 to 1.8, exceeding the Eshelby speed of $\sqrt{2} * V_s$ expected for rupture in a homogeneous medium (*53*). It also surpassed the predicted upper limit—the compliant side's compressional wave speed (V_p)—suggested by numerical simulations for the material contrast level across the Sagaing fault [$V_{s_compliant}/V_{s_stiff}$ = 0.85; (*46*, *50*)] (fig. S31). These observations imply that factors beyond the bimaterial effect may have also contributed to the extreme rupture acceleration. One potential contributor is the poroelastic effect. Along the bimaterial Sagaing fault, the compliant side consists of porous sedimentary rocks, whereas the stiffer side comprises less porous metamorphic rocks (*54*). This material contrast could enhance the bimaterial effect and further promote rupture acceleration (*55*). Additionally, an elevated stress level due to the long period of seismic quiescence as well as the free-surface effect (*39*, *56*) may have played roles. Near-fault investigations, including mapping of distributed deformation, rock sampling from boreholes, and cross-fault seismic reflection experiments, will help evaluate these possibilities.

The ratio between the dissipated and potential energies (G_c/G_0) is another key factor controlling rupture propagation (*57*). Here, G_c represents the fracture energy dissipated near the rupture front, whereas G_0 denotes the static elastic energy release rate for strike-slip rupture with a finite rupture width (*12*). Theory and numerical simulations have demonstrated that persistent supershear propagation is permitted when $G_c/G_0 < 0.7$ is satisfied on a purely strike-slip fault (*57*). The southern segment of the Sagaing fault, between Mandalay and NayPyiDaw, has remained seismically quiescent without experiencing any $M \geq 7$ earthquakes since 1839 (Fig. 1A) (*2*). This prolonged interseismic period results in considerable prestress accumulation and slip deficit. Because G_0 depends more strongly than G_c on slip deficit (*12*, *56*), the energy ratio G_c/G_0 decreases with longer interseismic intervals. Thus, the extended seismic quiescence led to another favorable condition for sustained supershear rupture.

Fault geometry also has an important influence on earthquake rupture propagation. We found that the southern segment of the Sagaing fault exhibits a remarkably simple and linear geometry, which is conducive to persistent supershear propagation (*58*). The three-segment fault model used in the joint FFI incorporated a total strike variation of only 9° (Fig. 2). From the epicenter to (21.51°N, 95.99°E), the first segment extends ~55 km with a strike of 179°. The following 110-km segment to the south adopts a strike of 176°, and then the strike shifts

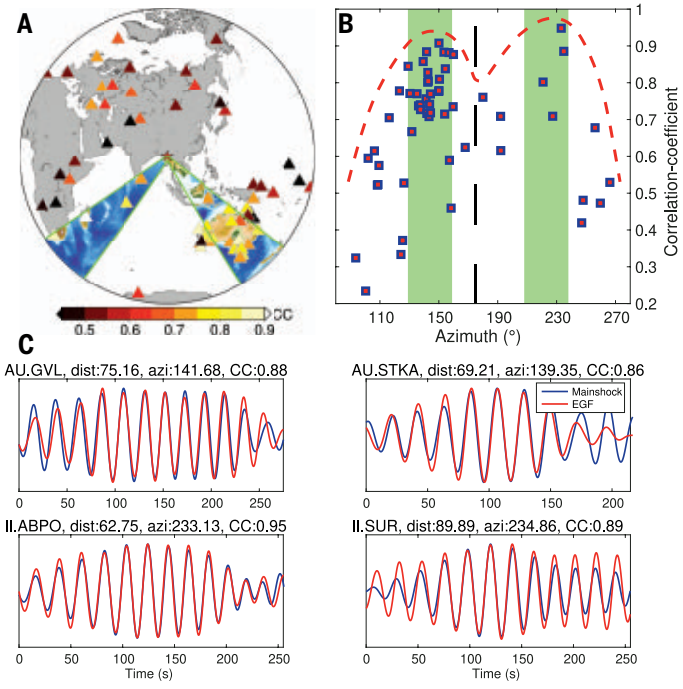


Fig. 4. Far-field Rayleigh wave Mach cones and waves. (**A**) Mach cone distribution. Colored areas within the green boundaries mark the Mach cones, where Mach waves have been identified. Seismic station locations are shown as triangles, with color indicating the CCs between 15- and 25-s Rayleigh wave displacement seismograms from the M_w 7.8 Mandalay mainshock and the M_w 5.1 EGF event. (**B**) Azimuthal distribution of waveform similarity. CCs between 15- and 25-s Rayleigh waveforms of the mainshock and EGF, plotted as a function of station azimuth. The black dashed line marks the rupture direction of the southern segment. Green bands denote the azimuthal extent of the identified Mach cones, accounting for uncertainties in Rayleigh wave phase and rupture velocities. The red dashed curve indicates the envelope that delineates the distribution pattern of CCs. (**C**) Representative Mach waves. Bandpass-filtered vertical displacement Rayleigh wave seismograms for the mainshock (blue) and EGF (red) at selected stations. Station name, azimuth (azi), and epicentral distance (dist; in degrees) are annotated.

to 172° for the next 220 km. Near the southern terminus, the strike rotates slightly to 170°. Whereas this three-segment model is a simplified representation of the fault based on ground deformation (Fig. 2B), detailed inspection of the surface trace indicates that the curve of the fault is continuous without any sizeable kinks or step-overs (fig. S32). Numerical simulations and laboratory experiments by (*59*, *60*) demonstrated that ruptures can propagate at supershear speeds across fault bends of 10°, whether on the extensional or compressional side, without any reduction in propagation velocity. Therefore, the geometrically simple and linear structure of the Sagaing fault likely helped to sustain supershear rupture propagation, consistent with observations in other large events (*58*).

REFERENCES AND NOTES

1. P. Bird, *Geochem. Geophys. Geosyst.* **4**, 2001GC000252 (2003).
2. Y. Wang, K. Sieh, S. T. Tun, K.-Y. Lai, T. Myint, *J. Geophys. Res. Solid Earth* **119**, 3767–3822 (2014).
3. G. Bertrand *et al.*, *C. R. Acad. Sci. II* **7**, 479–484 (1998).
4. C. Vigny *et al.*, *J. Geophys. Res.* **108** (B11), 2002JB001999 (2003).
5. X. Wang *et al.*, *J. Geophys. Res. Solid Earth* **124**, 504–526 (2019).
6. Z. Peng *et al.*, Mainshock rupture properties, aftershock activities and remotely triggered seismicity associated with the 2025 Mw 7.7 Sagaing Fault earthquake in Myanmar. EarthArXiv X5BQ7D (2025); https://doi.org/10.31223/X5BQ7D.
7. L. Meng, A. Zhang, Y. Yagi, *Geophys. Res. Lett.* **43**, 628–636 (2016).
8. C. Ji, D. J. Wald, D. V. Helmberger, *Bull. Seismol. Soc. Am.* **92**, 1192–1207 (2002).
9. G. Shao, X. Li, C. Ji, T. Maeda, *Earth Planets Space* **63**, 559–564 (2011).
10. S.-H. Yun *et al.*, *Seismol. Res. Lett.* **86**, 1549–1556 (2015).
11. Z. E. Ross *et al.*, *Science* **366**, 346–351 (2019).
12. Materials and methods are available as supplementary materials.
13. L. Xu *et al.*, *J. Geophys. Res. Solid Earth* **128**, e2022JB025936 (2023).
14. D. Melgar, G. P. Hayes, *Geophys. Res. Lett.* **44**, 9691–9698 (2017).
15. D. Wang, J. Mori, K. Koketsu, *Earth Planet. Sci. Lett.* **440**, 115–126 (2016).
16. L. Ye *et al.*, *Earth Planet. Sci. Lett.* **401**, 215–226 (2014).
17. L. Xu *et al.*, *Commun. Earth Environ.* **4**, 379 (2023).
18. L. T. Aung *et al.*, The 2012 Mw 6.8 Thabeikkyin earthquake in central Myanmar: A widely felt strong event with unilateral rupture through the Singu basalt on the Sagaing fault." in AGU Fall Meeting, 12 to 16 December 2022, Chicago, IL (AGU, 2022).
19. Y. Yang *et al.*, *Int. J. Appl. Earth Obs. Geoinf.* **133**, 104133 (2024).
20. S. Xu, J. Dimasaka, D. J. Wald, H. Y. Noh, *Nat. Commun.* **13**, 7793 (2022).
21. M. Vallée, E. M. Dunham, *Geophys. Res. Lett.* **39**, L05311 (2012).
22. H. Bao *et al.*, *Nat. Geosci.* **12**, 200–205 (2019).
23. M. Mello, H. S. Bhat, A. J. Rosakis, *J. Mech. Phys. Solids* **93**, 153–181 (2016).
24. V. Rubino, A. J. Rosakis, N. Lapusta, *J. Geophys. Res. Solid Earth* **125**, e2019JB018922 (2020).
25. E. M. Dunham, R. J. Archuleta, *Bull. Seismol. Soc. Am.* **94** (6B), S256–S268 (2004).
26. H. Bao *et al.*, *Nat. Geosci.* **15**, 942–949 (2022).
27. M. Abdelmeguid *et al.*, *Commun. Earth Environ.* **4**, 456 (2023).
28. C. Ren *et al.*, *Science* **383**, 305–311 (2024).
29. A. Rosakis, M. Abdelmeguid, A. Elbanna, *Nat. Geosci.* **18**, 534–541 (2025).
30. D. M. Boore, *Bull. Seismol. Soc. Am.* **91**, 1199–1211 (2001).
31. B. Delouis, M. van den Ende, J.-P. Ampuero, *Bull. Seismol. Soc. Am.* **114**, 726–740 (2023).
32. S. Yao, H. Yang, *Sci. Adv.* **11**, eadq0154 (2025).
33. H. Kanamori, E. E. Brodsky, *Rep. Prog. Phys.* **67**, 1429–1496 (2004).
34. R. Burridge, *Geophys. J. Int.* **35**, 439–455 (1973).
35. D. J. Andrews, *J. Geophys. Res.* **81**, 5679–5687 (1976).
36. Y. Liu, N. Lapusta, *J. Mech. Phys. Solids* **56**, 25–50 (2008).
37. E. M. Dunham, P. Favreau, J. M. Carlson, *Science* **299**, 1557–1559 (2003).
38. J. N. Hayek *et al.*, *Geophys. Res. Lett.* **51**, e2024GL110128 (2024).
39. Y. Kaneko, N. Lapusta, *Tectonophysics* **493**, 272–284 (2010).
40. L. Bruhat, Z. Fang, E. M. Dunham, *J. Geophys. Res. Solid Earth* **121**, 210–224 (2016).
41. X. Ma, A. Elbanna, *Earth Planet. Sci. Lett.* **523**, 115702 (2019).
42. K. J. Ryan, D. D. Oglesby, *J. Geophys. Res. Solid Earth* **119**, 5814–5829 (2014).
43. Y. Huang, J. P. Ampuero, D. V. Helmberger, *Earth Planet. Sci. Lett.* **433**, 109–115 (2016).
44. S. Langer, L. Olsen-Kettle, D. Weatherley, *Geophys. J. Int.* **190**, 1169–1180 (2012).
45. H. Shlomai, J. Fineberg, *Nat. Commun.* **7**, 11787 (2016).
46. H. Shlomai, M. Adda-Bedia, R. E. Arias, J. Fineberg, *J. Geophys. Res. Solid Earth* **125**, e2020JB019829 (2020).
47. A. Mitchell, The Shan Plateau West of the Salween, *Geological Belts, Plate Boundaries, and Mineral Deposits in Myanmar*, A. Mitchell, Ed. (Elsevier, 2018), pp. 19–76.
48. D. Boore, *Bull. Seismol. Soc. Am.* **106**, 316–320 (2016).
49. K. Xia, A. J. Rosakis, H. Kanamori, J. R. Rice, *Science* **308**, 681–684 (2005).
50. Z. Shi, Y. Ben-Zion, *Geophys. J. Int.* **165**, 469–484 (2006).
51. M. Abdelmeguid, A. Elbanna, *Earth Planet. Sci. Lett.* **593**, 117648 (2022).
52. D. J. Andrews, Y. Ben-Zion, *J. Geophys. Res.* **102** (B1), 553–571 (1997).
53. K. Xia, A. J. Rosakis, H. Kanamori, *Science* **303**, 1859–1861 (2004).
54. Z. Zhou *et al.*, *J. Oceanol. Limnol.* **42**, 1074–1086 (2024).
55. E. M. Dunham, J. R. Rice, *J. Geophys. Res.* **113** (B9), 2007JB005405 (2008).
56. J. Xu, H. Zhang, X. Chen, *Geophys. J. Int.* **202**, 2194–2206 (2015).
57. H. Weng, J. P. Ampuero, *Nat. Geosci.* **13**, 817–821 (2020).
58. M. Bouchon *et al.*, *Tectonophysics* **493**, 244–253 (2010).
59. C.-E. Rousseau, A. J. Rosakis, *J. Geophys. Res.* **108** (B9), 2002JB002310 (2003).
60. E. L. Templeton *et al.*, *J. Geophys. Res.* **114** (B8), 2008JB006174 (2009).
61. R. Burns, piyushrpt, Z. Yunjun, B. Marfito, C. Liang, H. Fattahi, E. Fielding, Microwave Remote Sensing Laboratory, G. Manipon, S. Mirzaee, S. Staniewicz, Andy CA6JAU, V. Schut, asolarte, mzzhong, dbekaert, S. S. Sangha, Russell, G. H. X. Shiroma, L. Zhu, shitong01, hbparache, J. H. Kennedy, F. Amelung, bbuzz31/isce2: Buzzanga et al., 2023, (v2.6.3), Zenodo (2023); https://doi.org/10.5281/zenodo.8157051.
62. L. Xu, Z. Yunjun, Y. Yang, Y. Wang, C. Hu, Geodetic Deformation, DPM products, and source imaging results in "Bimaterial Effect and Favorable Energy Ratio Enabled Supershear Rupture in the 2025 Mandalay Earthquake", Zenodo (2025); https://doi.org/10.5281/zenodo.17113746.
63. L. Xu, L. Meng, Codes for Slowness-Enhanced Back-Projection (SEBP), Zenodo (2024); https://doi.org/10.5281/zenodo.12801968.
64. U.S. Geological Survey, Earthquake Hazards Program, Advanced National Seismic System (ANSS) Comprehensive Catalog of Earthquake Events and Products: Various (2017); https://doi.org/10.5066/F7MS3QZH.
65. G. Ekström, M. Nettles, A. M. Dziewonski, *Phys. Earth Planet. Inter.* **200-201**, 1–9 (2012).

66. R. Newman *et al.*, Wilber 3: A Python-Django Web Application For Acquiring Large-scale Event-oriented Seismic Data, Incorporated Research Institutions for Seismology (2013); https://ds.iris.edu/wilber3/.
67. A. Strollo *et al.*, *Seismol. Res. Lett.* **92**, 1788–1795 (2021).
68. GEOFON Data Centre, GEOFON Seismic Network (GFZ Helmholtz Centre for Geosciences, 1993). https://doi.org/10.14470/TR560404.
69. M. Beyreuther *et al.*, *Seismol. Res. Lett.* **81**, 530–533 (2010).
70. WorldPop, WorldPop Global Project Population Data; https://hub.worldpop.org/.
71. U.S. Geological Survey, Earthquake Hazards Program; https://earthquake.usgs.gov/earthquakes/eventpage/us7000pn9s/ground-failure/about.
72. Changguang Satellite, Jilin-1 Repository; https://www.jl1mall.com/resrepo/.
73. D. Gansu, *RSC: Comparative Analysis Data Before and After the Myanmar Earthquake on March 28, 2025* (National Tibetan Plateau / Third Pole Environment Data Center, 2025).
74. P. Wessel, W. H. F. Smith, *Eos* **79**, 579 (1998).
75. R. B. Lohman, W. D. Barnhart, *J. Geophys. Res. Solid Earth* **115**, 12413 (2010).

ACKNOWLEDGMENTS

We thank reviewer M. Abdelmeguid and the two anonymous reviewers for their constructive comments. We thank China Earthquake Networks Center (CENC) for providing the aftershock catalog. We thank Y. Wang for providing the fault dataset of Myanmar. We thank the US Geological Survey for providing Sentinel-2 and surface rupture data. We thank the European Space Agency, China Centre for Resources Satellite Data and Application, and Aerospace Exploration Agency for providing the Sentinel-1, LuTan-1, and ALOS-2 SAR imagery. We thank Google Maps for providing topography and world street maps. **Funding:** This work is supported by the NSF grant EAR-2506181 (L.X., L.M.) and the National Natural Science Foundation of China grants 42388102 and 42174023 (W.X.). **Author contributions:** Conceptualization: L.M., C.J., Z.Y., L.X.; Methodology: L.X., C.J., L.M., Z.Y., Y.Y.; Investigation: L.X., Z.Y., Y.Y., Y.W., C.H., H.W., C.J., E.S.; Visualization: L.X., Z.Y., Y.Y., Y.W., C.H., W.X.; Supervision: L.M., C.J., Z.Y.; Writing – original draft: L.X., H.W., Z.Y., C.J., Y.Y.; Writing – review & editing: all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. The SAR processing is performed using the ISCE-2 software, available at (*61*). The geodetic dataset, DPM products, and source imaging results of this study are available at (*62*). The MATLAB code for SEBP is available at (*63*). The moment tensor solutions came from the USGS (http://earthquake.usgs.gov) (*64*) and the Global Centroid Moment Tensor project (http://www.globalcmt.org) (*65*). All teleseismic data can be downloaded through the IRIS Wilber 3 system (https://ds.iris.edu/wilber3/) (*66*) and ORFEUS (www.orfeus-eu.org) (*67*). The local strong motion data can be downloaded from IRIS Wilber 3 system (https://ds.iris.edu/wilber3/) (*66*) and GFZ Helmholtz Centre for Georesearch (https://geofon.gfz.de/) (*68*). The Python software package Obspy (www.obspy.org) (*69*) was used for seismic data requesting and waveform filtering. Some of the images in the figures were based on optical images from Google Map, World Hillshade, OpenStreetMap, Jilin-1, and GaoFen series of satellites. Population density data was from the WorldPop Global Project Population Data (https://hub.worldpop.org/) (*70*). The hazard model and fault traces data are available on USGS (https://earthquake.usgs.gov/earthquakes/eventpage/us7000pn9s/executive) (*71*). The Jilin-1 imagery source is available at https://www.jl1mall.com/resrepo/ (*72*). The Gaofen series datasets are provided by National Tibetan Plateau/Third Pole Environment Data Center (http://data.tpdc.ac.cn) (*73*). Figures were produced using Generic Mapping Tools (*74*), Obspy, and Matlab. The static displacement data was downsampled by InSamp (https://github.com/williamBarnhart/InSamp) (*75*). **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ady6100
Materials and Methods; Figs. S1 to S40; Tables S1 to S6; References (*76*–*118*); Movies S1 to S4

Submitted 28 April 2025; accepted 8 September 2025

10.1126/science.ady6100

PRIZE ESSAY

NEUROSCIENCE

Respecifying partners

Neural circuits are rewired by manipulating cell surface combinatorial code

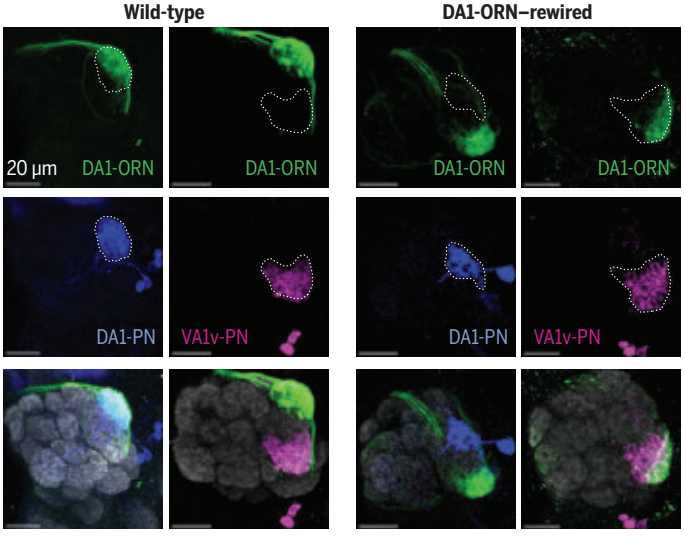
Cheng Lyu

Imagine being an axon during brain development. Your task is to find the correct synaptic partner in a brain filled with billions of neurons (*1*). It is a daunting challenge. The billion neurons in our brains form approximately 10^{14} synapses. That means that each axon must somehow identify the right partner from a vast number. How does the developing brain accomplish this feat of precision? And what happens to us when it fails?

One strategy that nature uses to make this complex problem more manageable is dimensionality reduction—essentially, it narrows down the search space. In some neural tissues, such as the layered vertebrate retina (*2*–*4*), the process is staged: First, neurons target a specific layer [a one-dimensional (1D) search], and then they look for the correct partner within that layer (a 2D search). But this simplification does not apply everywhere. For example, in adult fruit flies (*Drosophila melanogaster*), approximately 50 types of olfactory receptor neurons (ORNs) synapse with 50 types of projection neurons (PNs) in a one-to-one fashion at 50 discrete glomeruli, which together occupy stereotyped 3D positions in the antennal lobe. How is this feat achieved?

DA1-ORN axons rewired to VA1v-PN dendrites

Confocal images of four fly antennal lobes (each column shows the same lobe) showing DA1-ORN axons (top row), DA1-PN or VA1v-PN dendrites (middle row), and their merge with glomerular staining in gray (bottom row). PN dendrite borders are indicated with a dashed line. Wild-type flies (left two columns) and DA1-ORN–rewired flies (right two columns) are shown.



To answer this developmental question, it is necessary to study development dynamically, watching it unfold in real time. This is tricky: Gene expression in neurons changes rapidly as brains mature, which makes it difficult to track specific cell types using traditional genetic tools. To get around this hurdle, my colleagues and I developed a framework that allowed us to genetically label individual ORN and PN types consistently across developmental stages (*5*).

We discovered that although the adult antennal lobe is a 3D structure, the most critical phase of partner selection between ORNs and PNs occurs on a temporary 2D surface. During this period, ORN axons grow along the lobe surface, and PN dendrites extend outward to meet them—even those PNs that eventually settle deeper inside. This finding means that the initial partner search is effectively constrained to a 2D space.

The simplification does not stop there. We observed that each ORN type follows a genetically hardwired trajectory along the lobe surface, and this trajectory intersects with dendrites of its destined PN partner. This process further narrows the search from two dimensions to essentially one dimension. Each ORN "walks" a predetermined path and selects its partner from just a handful of candidate PN types along the way.

Although this model was largely built from observations of normal development across many ORN and PN types, it was also supported by cell type–specific perturbation experiments. When we genetically rerouted ORN axons away from their usual paths, their intended PN partners failed to migrate inward as usual. Instead, they remained stalled on the surface, as if waiting for their missing axonal partner. Moreover, the more that we rerouted the ORN trajectories, the more mismatched their connections became, which revealed a strong correlation between trajectory deviation and partner mismatch.

How does an ORN axon choose one synaptic partner out of a handful of potential candidates? Decades of research point to the role of cell surface proteins—ligands and receptors—that mediate attraction and repulsion between specific neuron types (*6*–*8*). But disrupting a single gene often yields only partial defects. Even null mutations, which completely eliminate gene function, typically produce only subtle or incomplete phenotypes. This result suggests that these proteins operate combinatorially or redundantly, which makes the system robust but also harder to decode using traditional "one-gene-at-a-time" methods.

An alternative approach to testing an understanding of a system is to rebuild it from key components. As the physicist Richard Feynman

GRAPHIC: ADAPTED FROM (*12*) BY A. FISHER/*SCIENCE*

once said, "What I cannot create, I do not understand." So we investigated whether we could rewire an ORN to form a new connection with a different PN by changing its cell surface molecular profile.

We focused on rewiring a group of ORNs that normally targets the DA1 glomerulus (DA1-ORNs) to instead form synapses with a group of PNs that normally receives input from the VA1v glomerulus (VA1v-PNs). We began by identifying candidate cell surface proteins using single-cell RNA sequencing (*9*, *10*) followed by genetic screening (*11*). We found dozens of proteins that, when manipulated, partially altered wiring patterns. We then performed systematic tests using RNA interference and gene overexpression tools to probe different combinations.

After extensive trials, we successfully rewired DA1-ORN axons to VA1v-PN dendrites by manipulating just five key proteins in one cell type (DA1-ORNs; see the figure), all in a fly engineered with 11 transgenes (*12*). We grouped our manipulations into three strategic categories: (i) increase repulsion between DA1-ORNs and their original DA1-PN partners, (ii) decrease repulsion between DA1-ORNs and the new VA1v-PN targets, and (iii) match attraction between DA1-ORNs and VA1v-PNs. No single manipulation resulted in major rewiring. A combination of two strategies produced intermediate effects. But when all three were applied, the rewiring was robust and consistent. Even more, this same strategy could be used to rewire other ORN-PN pairs, demonstrating the generalizability of the approach.

But do these anatomical changes translate to changes in function and behavior? To find out, we used calcium imaging to measure neuronal activity in live flies. We found that the rewired VA1v-PNs now responded to 11-*cis*-vaccenyl acetate (cVA), a pheromone typically only detected by DA1-ORNs (*13*, *14*). In wild-type flies, cVA detection suppresses male-male courtship (*14*). Meanwhile, odorants that activate VA1v-ORNs promote courtship (*15*). So, if VA1v-PNs now receive inputs from DA1-ORNs, the response to cVA might be misinterpreted or perhaps even reversed. We tested this idea by measuring courtship activity between males. As predicted, the rewired male flies appreciably increased courtship toward other males—a notable behavioral shift. This finding confirmed that altering neural wiring can indeed change how sensory information drives behavior.

Our study shows that synaptic specificity in developing brains can be determined by a small set of cell surface proteins acting in combination. These proteins orchestrate a combinatorial usage of attraction and repulsion that guides axons to their correct partners. By tweaking just a few of these components, we were able to respecify both neural and behavioral partners, opening doors for exploring how changes in circuit wiring contribute to neurodevelopmental disorders and how similar principles may have shaped the evolution of brain circuits across species. □

REFERENCES AND NOTES

1. F. A. C. Azevedo *et al.*, *J. Comp. Neurol.* **513**, 532 (2009).
2. J. R. Sanes, S. L. Zipursky, *Neuron* **66**, 15 (2010).
3. J. R. Sanes, S. L. Zipursky, *Cell* **181**, 536 (2020).
4. E. Agi *et al.*, *Science* **383**, 1084 (2024).
5. C. Lyu *et al.*, *Science* **388**, 538 (2025).
6. R. W. Sperry, *Proc. Natl. Acad. Sci. U.S.A.* **50**, 703 (1963).
7. B. J. Dickson, *Science* **298**, 1959 (2002).
8. A. L. Kolodkin, M. Tessier-Lavigne, *Cold Spring Harb. Perspect. Biol.* **3**, a001727 (2011).
9. C. N. McLaughlin *et al.*, *eLife* **10**, e63856 (2021).
10. Q. Xie *et al.*, *eLife* **10**, e63450 (2021).
11. Z. Li *et al.*, bioRxiv 10.1101/2025.03.01.640985 (2025).
12. C. Lyu, Z. Li, C. Xu, J. Kalai, L. Luo, bioRxiv 10.1101/2025.03.01.640986 (2025).
13. T. S. Ha, D. P. Smith, *J. Neurosci.* **26**, 8727 (2006).
14. A. Kurtovic, A. Widmer, B. J. Dickson, *Nature* **446**, 542 (2007).
15. H.-H. Lin *et al.*, *Neuron* **90**, 1272 (2016).

10.1126/science.aeb5128

Department of Biology, Stanford University, Stanford, CA, USA. Email: chenglyu@stanford.edu

PHOTOS: (TOP TO BOTTOM) COURTESY OF JUNJIE ZHANG; RYUJI SUZUKI; CLAIRE BALLARD

GRAND PRIZE WINNER



Cheng Lyu received an undergraduate degree from Peking University and a PhD from Rockefeller University. After completing his postdoctoral fellowship at Stanford University, he will start his laboratory in the School of Life Sciences at Westlake University in 2026. His research focuses on exploring how developmental changes in neural circuits shape animal behavior in ethologically meaningful ways. **www.science.org/doi/10.1126/science.aeb5128**

FINALISTS



Constanze Depp received an undergraduate degree from the Ruprecht Karl University of Heidelberg and a PhD from the Georg August University of Göttingen and the Max Planck Institute for Experimental Medicine. She is completing her postdoctoral fellowship in Beth Stevens's laboratory at Boston Children's Hospital and the Broad Institute of the Massachusetts Institute of Technology and Harvard University. Her research focuses on glia cell aging as a risk factor for Alzheimer's disease. **www.science.org/doi/10.1126/science.aeb5153**



Sara Mederos received an undergraduate degree from the Autonomous University of Madrid and a PhD from the Cajal Institute (CSIC) and the Complutense University of Madrid. She is currently a Wellcome Trust Early-Career Fellow at the Sainsbury Wellcome Centre (University College London) and will start her laboratory in the Neurosciences Programme at the Hospital del Mar Research Institute Barcelona in 2026. Her research investigates the brain circuits that enable flexible choices under changing internal states and external contexts. **www.science.org/doi/10.1126/science.aeb5165**

eppendorf & Science PRIZE FOR NEURO BIOLOGY

PRIZE ESSAY

NEUROSCIENCE

White matters arising

Rethinking Alzheimer's disease through the lens of myelin

Constanze Depp[1,2]

> *What's your name? – Auguste.*
> *Family name? – Auguste.*
> *What's your husband's name? – I think, Auguste."*

More than 100 years ago, German psychiatrist Alois Alzheimer noted down this dialogue in the medical records of his patient, Auguste Deter. At only 50 years of age, she presented with rapidly worsening signs of forgetfulness and disorientation. Today, this illness is well known as the most common form of dementia: Alzheimer's disease (AD). The condition largely remains a medical enigma even after a century's worth of research. Studies have long focused on the disease's neuropathological hallmarks, particularly protein aggregates composed of amyloid-β (Aβ) protein. These plaques are clustered in the brain's gray matter, where neuronal cell bodies reside. This lens has cast neurons as both victims and perpetrators of Aβ pathology. Yet neurons are not confined to the gray matter; their axonal projections stretch vast distances, with many traversing the brain's white matter highways.

The white matter derives its name and color from the lipid-rich myelin that wraps around axons, which enables rapid signal transmission. Once considered a static insulator, myelin is now known to be highly dynamic. Myelin is produced by oligodendrocytes in the central nervous system; it dynamically responds to neuronal activity and provides metabolic support to axonal segments far from their cell bodies (*1*). The aging brain undergoes progressive myelin deterioration, which is particularly pronounced in AD and is correlated with a more rapid cognitive decline (*2*). Despite critically assisting neurons, oligodendrocytes and myelin have largely been overlooked in AD pathogenesis. My colleagues and I wondered whether the disruption of myelin and the stress it exerts on neurons fuels Aβ deposition and AD pathogenesis as a whole. And if so, could this help explain why people become susceptible to AD with age?

The idea that myelin decline could play a role in AD pathologies was first brought forward by the late George Bartzokis in 2004 (*3*). However, this hypothesis was solely based on correlative human imaging studies and lacked experimental evidence. Inspired by this early work, we investigated whether myelin dysfunction could actively drive Aβ deposition in mouse models of AD. Indeed, when we chronically disturbed myelin function genetically, we discovered more Aβ deposits in the brains of these mice (*4*). Similarly, when myelin breakdown was induced acutely by a demyelinating agent, small Aβ aggregates were found around demyelinated lesions (*4*).

But how does myelin breakdown drive Aβ deposition? As myelin deteriorates, ensheathed axons suffer, get damaged, and often swell. We revealed that these axonal swellings accumulate all components necessary for Aβ production: Aβ precursor protein (APP) as well as β- and γ-secretases that cleave APP to release Aβ. Alongside altered Aβ production, we also investigated whether degradation of Aβ might be affected by myelin defects. Microglia—the brain's resident immune cells—phagocytose Aβ and cluster around plaques to protect surrounding tissue from toxic effects. However, in brains with both damaged myelin and Aβ plaques, microglia leave Aβ plaques unchecked because they are preoccupied with clearing myelin debris (*4*). Through single-cell RNA sequencing, we found that microglia exhibit a hybrid activation state: These multitasking microglia are essentially left overwhelmed because they are caught between clearing myelin debris and degrading Aβ (*4*). Defective myelin therefore drives plaque deposition in AD by a dual mechanism that involves both increased production and reduced clearance of Aβ.

The transgenic mouse model used in our study was engineered to drive Aβ expression exclusively in neurons to cater to a long-standing dogma in the field. However, our findings prompted us to also consider whether oligodendrocytes could contribute to amyloid burden directly by generating Aβ themselves. Indeed, single-cell transcriptomics revealed that oligodendrocytes express the full amyloidogenic machinery at levels comparable to or even higher than those in neurons (*5*, *6*). To assess the contribution of excitatory neurons and oligodendrocytes to Aβ plaque load, we turned to a new AD mouse model wherein all cell types that endogenously express the amyloidogenic machinery can, in principle, produce Aβ (*7*). We further modified these mice with cell type–specific deletion of the β-secretase gene *Bace1*, which produces the rate-limiting enzyme in Aβ production. When BACE1 production was ablated in oligodendrocytes, brainwide plaque burden decreased by 30%, which demonstrated that oligodendrocytes appreciably contribute to the Aβ pool (*5*). This observation was corroborated by two other independent studies that also used AD mice with BACE1-deficient oligodendrocytes (*8*, *9*). Deleting *Bace1* in excitatory neurons led to a 95% drop in plaque formation (*5*), which confirmed that excitatory neurons are the primary Aβ source for plaque formation. By teasing apart the cell type–specific contribution to Aβ plaque load, we observed that Aβ production and plaque deposition are nonlinearly related. Below a critical threshold, Aβ largely remains soluble and nonaggregated (*10*). Upon reaching this threshold, Aβ oligomers accumulate, and then fibril and plaque formation ensues. Though oligodendrocyte-derived Aβ alone cannot kickstart plaque aggregation, it exacerbates plaque deposition once neuronally derived Aβ reaches a tipping point. Our study indicates that initiating early interventions before Aβ levels cross this threshold could potentially prevent plaque formation altogether.

Together with other recent studies in the field, our observations invite a reexamination of AD with particular attention to direct and indirect glial contributions, especially from oligodendrocytes and myelin. Importantly, this view may inform new therapeutic strategies: Approaches to maintain myelin integrity throughout aging, such as remyelination-promoting drugs developed for multiple sclerosis, might delay AD onset by preserving axonal health and giving microglia a break (*11*). Targeting non-neuronal Aβ sources specifically also becomes compelling because although Aβ-processing enzymes seem essential for neuronal functions, they might be dispensable for myelin and oligodendrocyte function. Be it plaques or soluble forms of Aβ, even modest reductions of Aβ could lead to important benefits to brain function.

Our work has revealed that AD is more than just a disease of neurons—it affects the brain's entire delicate cellular ecosystem. Conserving this ecosystem requires that the AD research community adopt a holistic approach, considering all its components and their interplay to finally halt AD in this century. □

REFERENCES AND NOTES

1. K. A. Nave, *Nature* **468**, 244 (2010).
2. Z. Gong *et al.*, *Alzheimers Dement.* **19**, 3098 (2023).
3. G. Bartzokis, *Neurobiol. Aging* **25**, 5 (2004).
4. C. Depp *et al.*, *Nature* **618**, 349 (2023).

5. A. O. Sasmita *et al.*, *Nat. Neurosci.* **27**, 1668 (2024).
6. V. Gazestani *et al.*, *Cell* **186**, 4438 (2023).
7. T. Saito *et al.*, *Nat. Neurosci.* **17**, 661 (2014).
8. R. M. Rajani *et al.*, *PLOS Biol.* **22**, e3002727 (2024).
9. A. Ishii *et al.*, *Mol. Neurodegener.* **19**, 83 (2024).
10. J. Novo, S. Freire, W. Al-Soufi, *Sci. Rep.* **8**, 1783 (2018).
11. J.-F. Chen *et al.*, *Neuron* **109**, 2292 (2021).

10.1126/science.aeb5153

[1]Stevens Lab, Boston Children's Hospital, Boston, MA, USA. [2]Stanley Center for Psychiatric Research, Broad Institute, Cambridge, MA, USA. Email: depp@broadinstitute.org

FINALISTS



Constanze Depp received an undergraduate degree from the Ruprecht Karl University of Heidelberg and a PhD from the Georg August University of Göttingen and the Max Planck Institute for Experimental Medicine. She is completing her postdoctoral fellowship in Beth Stevens's laboratory at Boston Children's Hospital and the Broad Institute of the Massachusetts Institute of Technology and Harvard University. Her research focuses on glia cell aging as a risk factor for Alzheimer's disease. **www.science.org/doi/10.1126/science.aeb5153**

eppendorf & Science PRIZE FOR NEURO BIOLOGY

PHOTO: RYUJI SUZUKI

PRIZE ESSAY

NEUROSCIENCE

Fear protects. Until it doesn't.

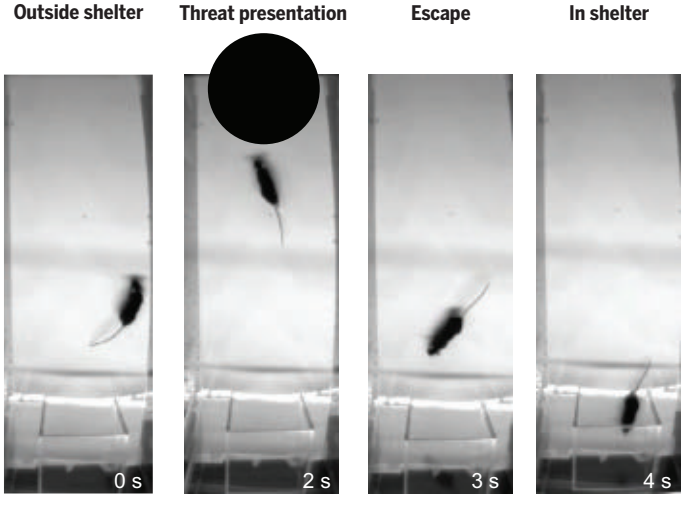
The brain learns to suppress instinctive defensive responses when threats prove harmless **Sara Mederos**

You stand in front of a room packed with listeners. It is quiet. Everything inside you tenses up. Slides flicker onto the screen. Suddenly you are aware of your breath, your heartbeat, the stillness. Thoughts you had perfectly lined up begin to scatter. Will your voice shake? Will someone ask something you cannot answer? You tell yourself it is just a talk. But your body feels otherwise. Where our ancestors ran from a wild bear in the forest, in our modern society, we face pressure, deadlines, the fear of failing.

Eventually, with experience, it starts to feel natural and perhaps even fun. The nerves are still there, but they become familiar. You are no longer ruled by fear. You figure out when to breathe. Although physical threats such as predators are rare in modern life, the brain continues to respond to social and professional stressors with the same ancient machinery (*1*). Situations involving evaluation, uncertainty, or perceived loss of control, such as public speaking, job interviews, or social exclusion, can activate defensive circuits originally shaped to manage life-or-death encounters. When these responses fail to adjust, the same systems that once kept us safe can become overreactive and disrupt daily life. Chronic stress in contemporary environments has been shown to engage circuits evolved for acute survival and contribute to the persistence of anxiety and maladaptive fear with a high prevalence in industrialized societies (*2*–*4*). The amygdala, hypothalamus, and periaqueductal gray, all central to instinctive defensive behaviors, remain key players in how we process these modern challenges (*1*, *5*). Yet, while many of the brain systems that regulate these behaviors are the same, knowledge remains lacking about how the brain learns to adapt defensive responses and suppress them when there is no reason for fear.

Snapshots of escape behavior evoked by a visual threat

Representative video frames showing a mouse exploring an open arena and initiating a rapid escape to a shelter after a looming stimulus presentation (from left to right), illustrating the behavioral response to a perceived threat. Time of the behavioral sequence in seconds is shown at the bottom right of each frame.



To try to understand this process, we used a visual threat paradigm in mice—an expanding dark disk presented from above—to mimic an approaching predator (*6*). Mice instinctively respond to this stimulus by rapidly escaping to a shelter, the safest place in the environment. However, when this visual threat is repeated without negative consequences, the animals' behavior gradually shifts as they learn that this stimulus is harmless, until they completely suppress these responses (*7*), which allows them to explore, forage, or seek a mate (*8*). This experiment enabled us to ask how the brain learns to suppress this instinctive response by mapping the circuits that could be involved in this behavioral adaption.

My colleagues and I focused on higher areas of the visual cortex, which processes complex visual information. These higher visual areas (HVAs) flexibly integrate sensory input with internal states such as locomotion and motivation, to allow visual processing to adapt dynamically to behavioral context and task demands (*9*). With optogenetics, we could specifically silence these areas as the mice were confronted with the simulated predator. We found that the HVAs were not necessary for the escape itself, yet they were involved in learning. When we interfered with their activity during repeated exposure to the stimulus, the mice continued to escape and did not learn to suppress this behavior. However, once they had learned, HVAs were no longer necessary, which suggested that the memory had been transferred elsewhere.

We traced the HVA outputs to the ventrolateral geniculate nucleus (vLGN), a small structure in the prethalamus full of inhibitory neurons that has been described to have strong control over areas involved in the threat escape response (*10*, *11*). Mice that had learned to suppress escape responses fled to shelter once again when the vLGN was silenced. Additionally, activation of the vLGN cells that received input from the HVAs in naïve animals was enough to suppress the escape response even without any learning. The vLGN had taken over as the point of control of the responses once the animals had learned them. We recorded vLGN activity with silicon probes as animals underwent learning to suppress escape responses, to identify neural signatures of this process. We found that specifically inhibitory vLGN neurons that received input from HVAs increased their responses to the stimuli, which allowed the mice to suppress escape. This response increase was due to a specific type of plasticity in vLGN, a reduction in inhibition onto these neurons, known as inhibitory long-term depression. This process was mediated by endocannabinoid signaling, which involved brain-internal retrograde messenger molecules known to regulate mood and memory (*12*, *13*) released by postsynaptic cells in the vLGN. Consequently, animals failed to learn to suppress escape responses when we pharmacologically blocked endocannabinoid release or their receptors. To test this form of plasticity more directly, we turned to acute brain slice recordings. We targeted the population of vLGN neurons that received input from HVAs and compared their activity in naïve animals and those that had undergone learning. The results corroborated our in vivo observations: Cells from animals that had undergone learning showed less inhibitory input, which confirmed a disinhibitory mechanism that reduced presynaptic γ-aminobutyric acid release onto these cells and allowed this population of inhibitory neurons to become more active and suppress the escape behavior.

GRAPHIC: ADAPTED BY A. FISHER/*SCIENCE*

We described a mechanism by which the brain adapts defensive responses on the basis of experience (*14*). It has long been suggested that one of the roles of the sensory cortex is not just to represent the outside world but to shape how we respond to it, to modify the intensity of instinctive behaviors through its influence on subcortical circuits, and our study found direct evidence for this function. By understanding how the brain learns not just to fear but to let defensive responses go, we could also shed light on dysregulation. In a world where anxiety often stems from past threats that no longer exist, understanding adaptation may be the key. In conditions such as phobias, anxiety, or post-traumatic stress disorder (PTSD), defensive responses can persist when there is no real threat (*15*), which suggests that defensive-suppressing circuits may be impaired. By understanding how this works at the level of specific brain regions and signaling systems, such as the endocannabinoids in the vLGN, we open a window that eventually could lead to new ways of thinking about and treating anxiety disorders and PTSD. □

REFERENCES AND NOTES

1. V. Taschereau-Dumouchel, M. Michel, H. Lau, S. G. Hofmann, J. E. LeDoux, *Mol. Psychiatry* **27**, 1322 (2022).
2. B. S. McEwen, *Physiol. Rev.* **87**, 873 (2007).
3. E. R. Duval, A. Javanbakht, I. Liberzon, *Ther. Clin. Risk Manag.* **11**, 115 (2015).
4. B. S. McEwen, C. Nasca, J. D. Gray, *Neuropsychopharmacology* **41**, 3 (2016).
5. P. Tovote, J. P. Fadok, A. Lüthi, *Nat. Rev. Neurosci.* **16**, 317 (2015).
6. M. Yilmaz, M. Meister, *Curr. Biol.* **23**, 2011 (2013).
7. S. C. Lenzi *et al.*, *Curr. Biol.* **32**, 2972 (2022).
8. D. A. Evans, A. V. Stempel, R. Vale, T. Branco, *Trends Cogn. Sci.* **23**, 334 (2019).
9. L. L. Glickfeld, S. R. Olsen, *Annu. Rev. Vis. Sci.* **3**, 251 (2017).
10. L. D. Salay, A. D. Huberman, *Cell Rep.* **37**, 109792 (2021
11. A. Fratzl, S. B. Hofer, *Neuron* **110**, 2728 (2022).
12. V. Chevaleyre, K. A. Takahashi, P. E. Castillo, *Annu. Rev. Neurosci.* **29**, 37 (2006).
13. B. D. Heifets, P. E. Castillo, *Annu. Rev. Physiol.* **71**, 283 (2009).
14. S. Mederos, P. Blakely, N. Vissers, C. Clopath, S. B. Hofer, *Science* **387**, 682 (2025).
15. A. Etkin, T. D. Wager, *Am. J. Psychiatry* **164**, 1476 (2007).

10.1126/science.aeb5165

Sainsbury Wellcome Centre for Neural Circuits and Behaviour, London, UK.
Email: smederos.cr@gmail.com

FINALISTS



Sara Mederos received an undergraduate degree from the Autonomous University of Madrid and a PhD from the Cajal Institute (CSIC) and the Complutense University of Madrid. She is currently a Wellcome Trust Early-Career Fellow at the Sainsbury Wellcome Centre (University College London) and will start her laboratory in the Neurosciences Programme at the Hospital del Mar Research Institute Barcelona in 2026. Her research investigates the brain circuits that enable flexible choices under changing internal states and external contexts.
www.science.org/doi/10.1126/science.aeb5165

eppendorf & Science PRIZE FOR NEURO BIOLOGY

PHOTO: CLAIRE BALLARD

REVIEW SUMMARY

MATERIALS SCIENCE

Porous materials: The next frontier in energy technologies

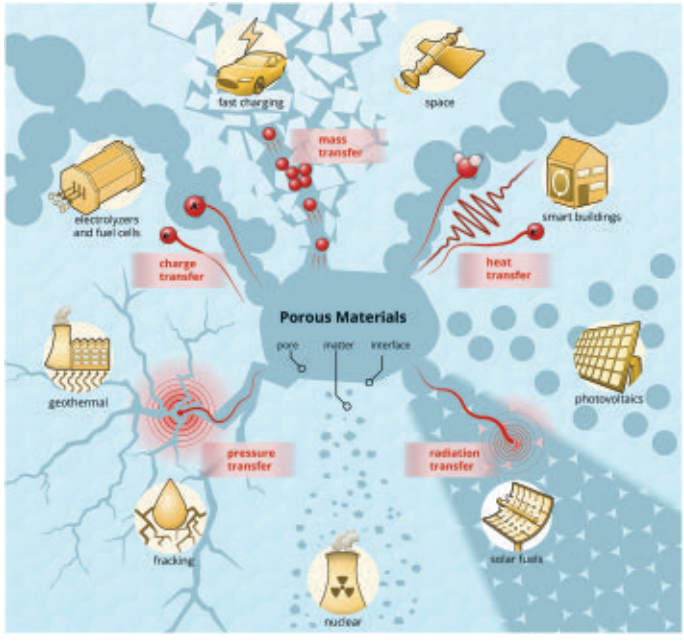
Eliyahu M. Farber, Nicola M. Seraphim, Kesha Tamakuwala, Andreas Stein, Maja Rücker, David Eisenberg*

Full article and list of author affiliations: https://doi.org/10.1126/science.adn9391

BACKGROUND: Porous materials, with pores ranging from angstroms to centimeters, consist of two phases—void and matter. These phases allow the transfer of multiple energy vectors, such as mass, charge, heat, radiation, and pressure. The high-area interface between the phases allows for these energy vectors to meet and interconvert by various processes, ranging from desired catalytic conversions between light, heat, electricity, and chemical streams to undesired events, such as cracking, passivation, or quenching of excited states. Thanks to their energy transfer capabilities, porous materials are increasingly adopted in a broad range of energy applications, driving performance breakthroughs in solar, nuclear, electrochemical, thermal, and subsurface energy extraction and conversion.

ADVANCES: By considering energy conversion devices or fuel extraction techniques in terms of energy vectors being simultaneously transferred and interconverted in the solid and void phases of a porous material, it is possible to design better porous components for a variety of applications. Such a view highlights the constraints and reveals new opportunities for optimization, simulation, and cross-disciplinary fertilization. For instance, the simultaneous tuning of mass and charge transfer in porous electrodes where the total volumetric surface area has been maximized leads to higher power density in electrochemical devices, such as supercapacitors, fuel cells, batteries, and electrolyzers. All energy streams are constrained by the geometry (size, shape, and orientation) and spatial distribution (ordering, hierarchy, and directionality) of the pores. Thus, heat storage in pore-embedded phase change materials (such as paraffin in carbons) is boosted by tuning pore sizes and apertures—improving heat conductivity while retaining the active matter. In thermochemically reactive materials that store solar heat at high temperatures, the porous structure can be tuned across multiple length-scales, enhancing the simultaneous absorption of light, distribution of heat, and evolution of the reaction front in time and space. Pores can also be used for selective blocking of energy streams: In thermoelectric generators, disconnected pores are introduced to disperse phonons but not electrons, reducing heat conductivity while maintaining the temperature difference and electronic conductivity needed for efficient thermoelectric energy conversion. Porous structures can evolve over time: Pores are formed inside nuclear fuel pellets by the action of evolved gases; by nanometric cracking of battery particles; or in the soil during unconventional extraction of oil, gas, and geothermal heat from low-porosity rocks (such as hydraulic fracturing).

OUTLOOK: The ultimate goal in porosity design—a task-specific porous architecture for each application—calls for great advances in analysis, synthesis, and material characterization. Better mathematical and computational models are needed for studying energy transfer processes in porous materials. These, in turn, require the complete three-dimensional visualization of porous structures on all length-scales, using tomographic, spectroscopic, and diffraction techniques. These will guide synthetic advances, leading to fuller control of the distribution of pore sizes, shapes, and connectivity patterns in materials. By taking a broader view of porous materials, commonalities between porosity subfields must be recognized, leading to the exchange of synthetic strategies, analytical tools, and inspiration sources. For example, biomimetically inspired design strategies could tackle challenges such as the simultaneous control of the internal (e.g., hierarchical) porosity of a material and its external morphology (determining textural porosity and roughness) or the synthesis of precise pore architectures without wasteful sacrificial templates. Tomographic and computational methods used for understanding fluid flow in subsurface energy extraction should be embraced even more by researchers in the fields of catalysis or heat storage. Overall, a deeper understanding of energy streams will drive new generations of porous components for our expanding energy landscape. □



Porous materials are key components of most energy conversion and extraction technologies. With pores ranging from angstroms to centimeters, porous materials can conduct the transfer of mass, charge, heat, radiation, and pressure, enabling energy conversion across many length-, time-, and energy scales.

*Corresponding author. Email: eisenberg@technion.ac.il Cite this article as E. M. Farber *et al.*, *Science* **390**, eadn9391 (2025). DOI: 10.1126/science.adn9391

MATERIALS SCIENCE

Porous materials: The next frontier in energy technologies

Eliyahu M. Farber[1], Nicola M. Seraphim[1], Kesha Tamakuwala[1]†, Andreas Stein[2], Maja Rücker[3,4], David Eisenberg[1]*

Porous materials with pore sizes spanning the range from molecular to macroscopic dimensions (from angstroms to centimeters) are essential in electrochemical, thermoelectric, nuclear, and solar power sources and in the extraction of oil, gas, and geothermal heat. To enable the clean, fast, and efficient conversion of energy, the porous structure must be designed to allow, modulate, or block the flow of energy transfer vectors. The most important energy streams are mass, charge, heat, radiation, and pressure, and they must be optimized while packing the optimal surface area per device volume. In this Review, we analyze the physical processes that enable energy transfer in porous structures, highlighting recent advances in the design, characterization, modeling, and fundamental understanding of porosity that have enabled breakthroughs across the landscape of energy technologies.

A porous material is composed of two phases—matter and void—and their various interfaces (Fig. 1A). Each phase can transport different energy carriers (molecules, ions, electrons, phonons, or photons), and the large interfacial area allows packing many active sites in a given volume. Porosity—understood broadly to mean the state of being porous, although the term denotes the void fraction (ε)—thus governs energy conversion in devices (such as porous electrodes or light absorbers) and media from which energy carriers are extracted (e.g., oil or hot fluid from porous rocks) (*1*). The connectivity of the void phase is defined by the pore diameters (d_{pore}) and their relative distances (d_{inter}), as depicted by the "connectivity diamond" (Fig. 1B). In some applications, open pores are needed to enhance mass transfer, whereas closed pores are preferred in others for blocking unwanted heat carriers. Different classes of porous structures can range from textural (formed naturally by the packing of solid particles) to designed structures (Fig. 1C), differing in the pore periodicity, directionality, and size distribution (from angstroms to centimeters; Fig. 1D). In this Review, we analyze the fundamental energy transport processes (Fig. 1E) that drive the adoption of porous materials in most energy conversion and extraction technologies, from oil recovery to thermoelectric propulsion, from solar fuels to nuclear power plants, and from batteries to fuel cells. Inspired to seek unity in variety and purpose in progress (*2*), we first discuss surfaces and interfaces in porous solids and then the transfer and storage of mass, charge, heat, radiation, and pressure. We conclude with brief comments on the future of porosity research in energy technologies.

Mass and pressure transfer in the porous subsurface

Mass transfer is perhaps the most defining property of porosity, even sharing its name: πορος in Greek means "passage," so "transport" is "passage across" (in this work, we use the term "transfer"). Subsurface mass transfer occurs across many length-scales, from nanometric pores to kilometer-wide caves. The sedimentary rock in conventional reservoirs, such as in the Middle East or the North Sea, contains textural porosity (arising from the packing and transformations of soil and inorganic particles) permeated by alkane-rich hydrocarbons. The porosity ε of fuel-rich sandstones and carbonate rocks is 0.2 to 0.3 and the permeability k is 10^{-4} to 10^{-2} darcy, whereas dense rock (such as clay, shale, or granite; $\varepsilon < 0.04$, $k = 10^{-9}$ darcy) must be fractured to yield the fuel or the geothermal heat.

Fuel extraction from conventional reservoirs involves single- and multiphase flow of fluids, such as hydrocarbons and injected brine (Fig. 2A). The pore structure of the rock determines K, the hydraulic conductivity in the flow direction, as well as the fluid front stability (along with fluid viscosity μ and surface tension), which is crucial for efficient recovery (Fig. 2B). The flow in porous rock is described by Darcy's law for steady, incompressible Newtonian fluids, $q = -\frac{K}{\mu}\left(\frac{dP}{dx}\right)$, where q is the instantaneous flux (volume per cross-sectional area) and $\frac{dP}{dx}$ the pressure gradient. The flow can be described numerically using the Brinkman extension for Darcy's law, if average permeability is assigned to regions with smaller pores (*3*). This approach can also account for Knudsen diffusion, where fluid molecules collide with walls more often than with each other ($d_{pore} < \sim 50$ nm), which leads to reduced effective diffusivity and surface-dominated transport.

The pore structure determines both the overall permeability of the rock and the local, complex, and sudden dynamics of fluid pockets (ganglia). The fluxes of both oil and brine are determined by their relative permeabilities through the rock ($k_{ri} = 0$ to 1) and their relative volume fractions (saturations) through a phenomenological extension of Darcy's law, $q_i = -\frac{k_{ri}K}{\mu_i}\left(\frac{dP_i}{dx} - \rho_i g\right)$, where ρ_i is fluid density and g is gravitational acceleration (*4*). Although this model assumes a stable flow pattern for each fluid, multiphase flow in porous rocks may also change rapidly because flow paths can break and reconnect through a complex interplay between viscous and capillary forces (*5*–*7*). The capillary pressure at pore throats (P_c) determines this complex flow, along with fluid properties (composition and saturation history) and rock conditions (pressure, temperature, and surface chemistry). This is described by the Young-Laplace equation, $P_c = \frac{4\gamma\cos\theta}{d_{throat}}$, where γ is the interfacial tension between the two fluids and θ is the local contact angle, expressing "wettability"—the local interplay between the surface energies of the three phases. The buildup of capillary pressure in pore throats causes sudden events, such as snap-off or coalescence of oil droplets (ganglions), piston-like filling of pores (*8*), or Haines jumps (*9*), a cascade of sudden pore-filling events upon pressure release (Fig. 2B) (*10*). The Young-Laplace equation also describes electrolyte wetting in the microporous electrodes of electrochemical power sources (*11*).

Recent advances in analytical techniques allow for the estimation of hydraulic conductivity from two-dimensional (2D) and 3D imaging of core sections, using electron microscopy, nano– and micro–computed tomography, magnetic resonance, and atomic force microscopy, and then analyzing the pore structure computationally using pore network models (*12*–*14*) (Fig. 2C). The 3D structure of fluids in pores can be analyzed geometrically using the four Minkowski functionals (*15*)—volume, surface area, mean curvature, and Gaussian curvature (reflecting local pore connectivity) (*16*, *17*). Eventually, macroscopic flow observations can be related to fundamental principles in thermodynamics and statistical mechanics (*18*, *19*).

The energy extracted from the subsurface can be enhanced at least fivefold (*20*) with enhanced (unconventional) extraction, where unconnected pores are accessed by hydraulic fracturing ("fracking"). The injection of high-pressure fluid props open fracture networks in the rock (*21*) (Fig. 2B), bridging nanometer-sized pores with meter-scale fractures, raising the porosity by orders of magnitude (ε = 0.1 to 0.4)

[1]Schulich Faculty of Chemistry, Grand Technion Energy Program, and Resnick Sustainability Center for Catalysis, Technion–Israel Institute of Technology, Haifa, Israel. [2]Department of Chemistry, University of Minnesota, Minneapolis, MN, USA. [3]Department of Mechanical Engineering, Eindhoven University of Technology, and Eindhoven Institute for Renewable Energy Systems, Eindhoven, Netherlands. [4]Max Planck Institute for Polymer Research, Mainz, Germany. ***Corresponding author. Email: eisenberg@technion.ac.il** †Present address: Chemistry Department, Stanford University, Stanford, CA, USA.

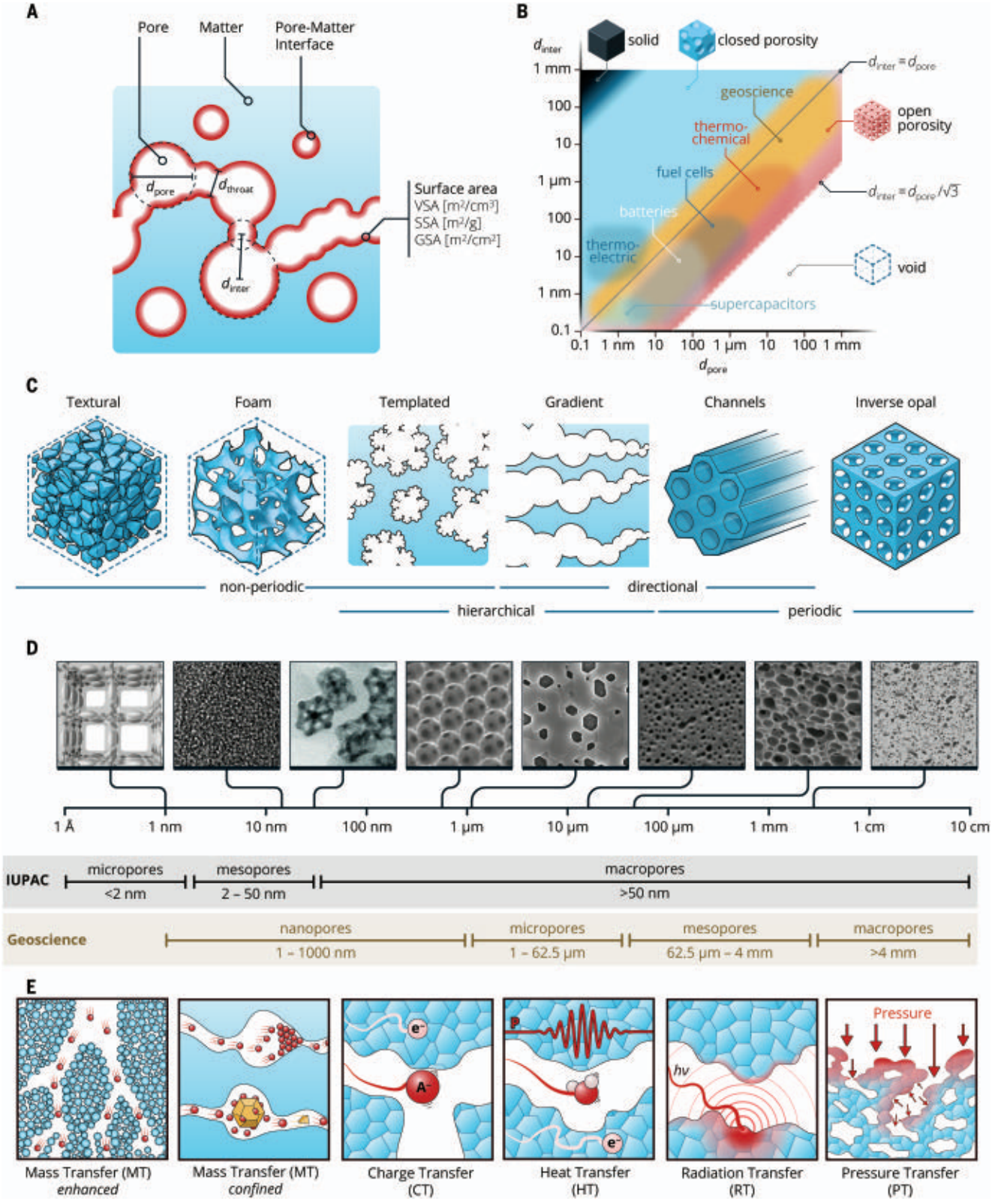


Fig. 1. Porous materials are central to energy technologies because both the void and solid phases conduct multiple energy transport processes. (**A**) A porous material is a two-phase composite of matter and void. The surface area (SA) can be normalized relative to the device or material volume (VSA), its specific mass (SSA), or its projected geometry (GSA). (**B**) The connectivity diamond: Pore diameters and distances determine the applications. Mass transfer is free in open and interconnected porosity, when $1 < \frac{d_{inter}}{d_{pore}} < \sqrt{3}$ (in 3D; assuming identical, equally spaced pores). (**C**) Classes of porous materials. (**D**) Micrographs or simulations of pores found in energy technologies ranging from supercapacitors to subsurface energy (reproduced with permission from Springer-Nature, RSC, AAAS, Wiley, and ACS). Pore nomenclatures of the International Union of Pure and Applied Chemistry (IUPAC) (*167*) and geoscience (*168*) are used. (**E**) Fundamental transfer phenomena governing energy conversion and extraction in porous materials: Each phase can allow, modulate, or block the transfer of mass, charge, heat, radiation, and pressure.

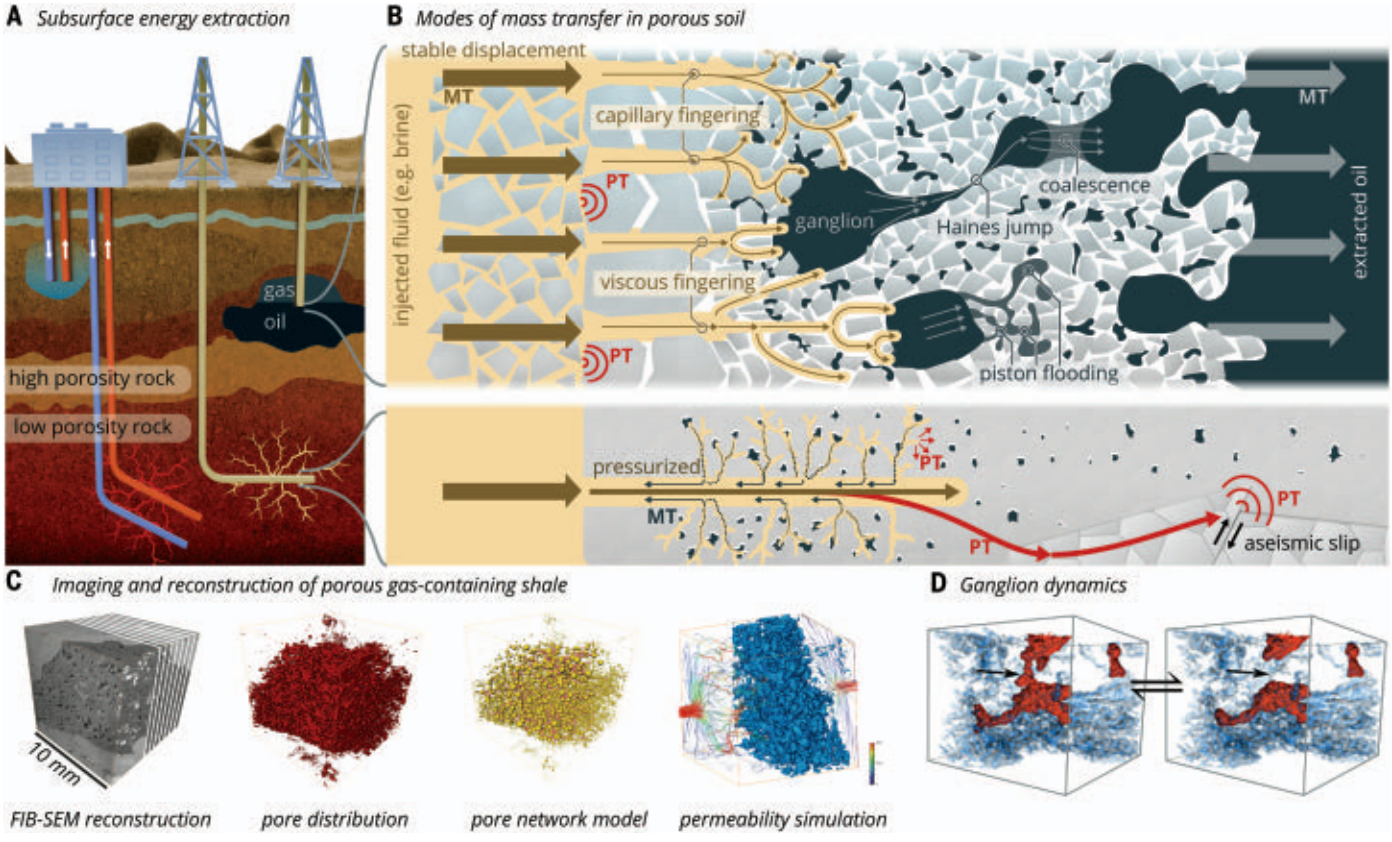


Fig. 2. Energy extraction from porous subsurface reservoirs. (**A**) Extraction of oil, gas, or geothermal heat, with or without hydraulic fracturing of the rock. (**B**) Modes of mass transfer: Stable viscous flow of injected and extracted fluids (such as brine and oil or cold and hot water in geothermal extraction), in parallel with complex dynamics of fluid pockets (ganglions) inside the pores, including corner flow due to capillary forces, Haines jump leading to ganglion coalescence, or piston-like flooding. MT, mass transfer; PT, pressure transfer. (**C**) Focused ion beam–scanning electron microscopy (FIB-SEM) imaging and 3D reconstruction of a shale core, allowing for pore network analysis and permeability simulations (*169*). (**D**) Micro–computed tomography imaging of ganglion dynamics in multiphase flow (*170*).

and increasing the effective drainage volume of the well. The optimal fracture network is hierarchical to hasten the fuel release ($d_{crack} = 10^{-6}$ to 10^{-3} m). Fracking is also useful for geothermal heat extraction, distributing the cold water in the hot underground rock (*22*).

The change in fluid pressure may induce local seismicity by triggering instabilities in prestressed faults or activate aseismic slip, outpacing the fluid migration (Fig. 2B). To understand pressure transfer in porous media (*23*, *24*), multiscale approaches join to describe multiphase fluid flow, the geomechanics of propagating fractures (*25*), thermal conductivity (in geothermal heat), electrical conductivity (for soil diagnostics), and even biochemistry (for pore clogging by soil microbes). Fluid pressure also causes pore deformation, coalescence, and cracking (*26*) in electrochemical (*27*) and thermal devices (*28*). Thus, models relating microscopic and heterogeneous features with the formation and propagations of large-scale transfer events are also promising for understanding electrochemical devices, such as fuel cells and electrolyzers, where pressure, heat, and charge transfer are combined with complex fluid flow and confinement dynamics (*29*, *30*).

Surfaces and interfaces in electrochemical devices

Supercapacitors, or electrochemical double-layer capacitors, store electrical energy by electrostatic adsorption of ions at the electrode surface with a given area (A), delivering high power with high fidelity (>10^6 cycles) at voltages (U) up to 4 V (*31*, *32*). During operation, electrons (flowing through the solid) and oppositely charged ions (flowing through the pores) meet at the interface, converting their kinetic electrical energies to potential electrical energy (E), which scales with the cell capacitance (C_{cell}), $E = \frac{1}{2} C_{cell} U^2$. C_{cell}, in turn, depends on the capacitance of each porous electrode, derived from its surface area, $C_{elec} = \epsilon A / d$ (where ϵ is the relative dielectric permittivity, A is the area of the interface, and d is the separation between charges). In typical electrodes, porous carbons are filled with aqueous, organic, or ionic liquid electrolytes (*33*). Micropores (<2 nm) contribute to high specific surface areas (up to 3000 m^2/g) and 2- to 4-nm mesopores boost mass transfer in an electronically conductive and chemically tunable matrix (*34*). Ultramicropores (d_{pore} < 1 nm), formerly considered closed for ion adsorption, may boost the capacitance with similarly sized electrolyte ions (*35*, *36*). This has sparked a debate about capacitance mechanisms, including on the meaning and measurement of surface area in a solid capable of intercalation between its layers (*37*). Thus, adjusting the pore size for optimal ion confinement while maximizing volumetric surface area can increase capacitance (Fig. 3B) (*38*). Such ion confinement also affects yield and selectivity in catalysis by prolonging the retention time of intermediates in catalyst pores. Similarly, the entropy loss arising from anion confinement in metal-organic frameworks (MOFs) boosts energy storage in a redox flow battery (*39*).

Supercapacitor electrodes can also store energy in the continuum between fast redox reactions and intercalation ("pseudocapacitance") (*40*), boosting the energy density by 10- to 100-fold at the expense of power density (*41*). Although pseudocapacitance is possible in carbons by surface functionalization (*42*), it is most pronounced in metal oxides (*43*) and sulfides (*43*). In MXenes, layered transition metal carbides, or carbonitrides (*44*), porosity is tuned by foaming (*45*), varying interlayer

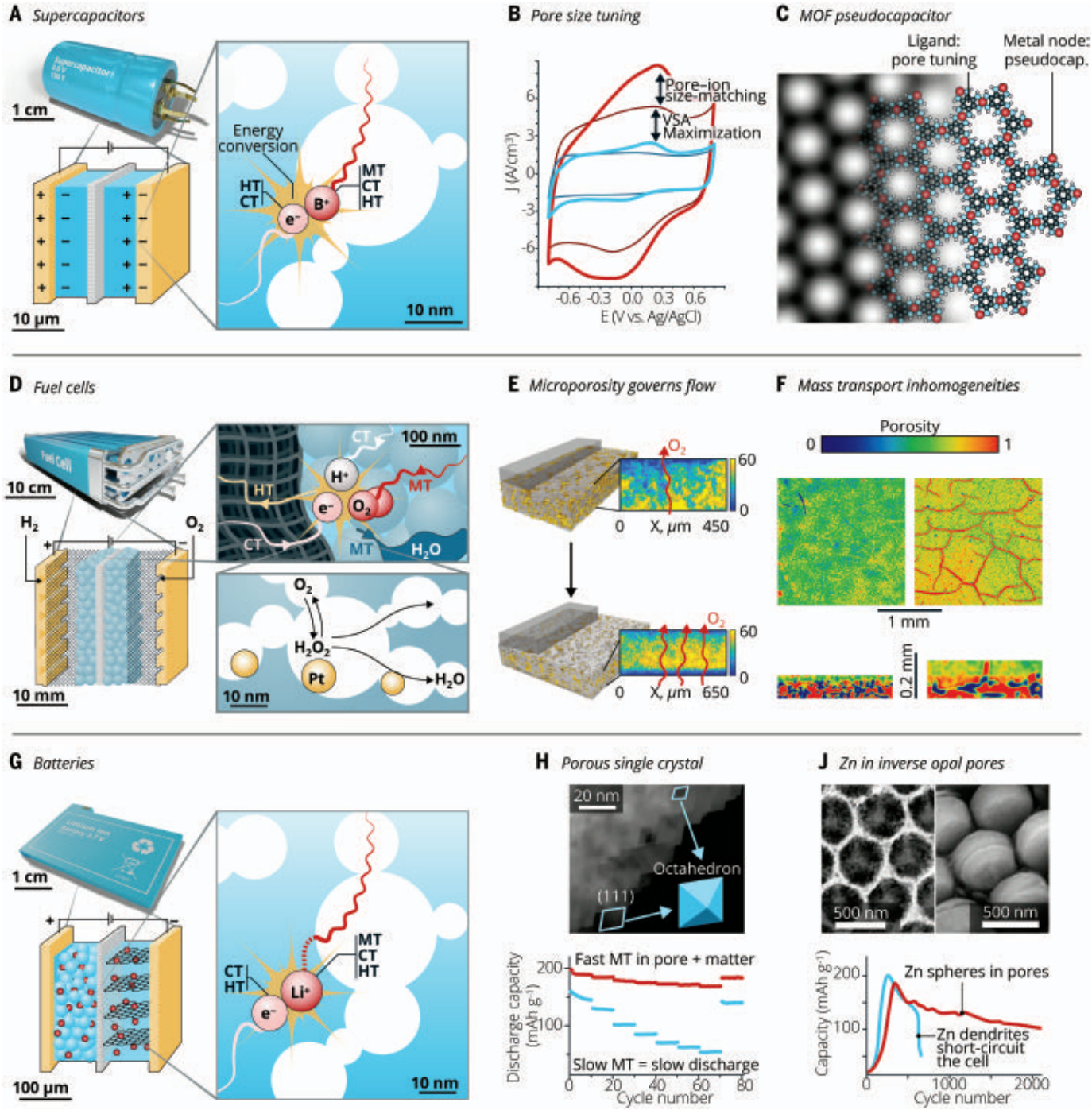


Fig. 3. Electrochemical energy conversion by porous materials in supercapacitors, fuel cells, and Li-ion batteries. (**A** to **C**) Supercapacitors. (**D** to **F**) Fuel cells. (**G** to **I**) Li-ion batteries. The operational principles point to the transport phenomena occurring in each application, including mass transfer (MT), charge transfer (CT), and heat transfer (HT). (B) Capacitance is increased by packing more surface area per volume and matching pore size to ion size for optimal confinement (*38*). (C) Tuning a 2D MOF allows packing more pores per volume, and the metal node provides pseudocapacitive energy storage (*49*). (E) A thickness-tuned MPL switches the O_2 flow mode from capillary fingering to viscous fingering (*171*). (F) X-ray tomography reveals preferential mass transfer channels in inhomogeneous gas transport layers (*56*). (H) Fast Li^+ transfer through nanometric pores in a lithium titanate anode, coupled with fast solid-state transfer through the boundary-free single-crystalline matrix, preserves battery capacity even at high discharge rates (*79*). (J) Zn dendrites are prevented in inverse opal porous electrodes due to confined Zn deposition and a homogeneous electric field (*81*).

spacing (*46*), or introducing in-plane pores (*47*). MOFs offer huge surface areas (up to 6000 to 8000 m^2/g), precisely tunable pores by ligand tailoring, and rich surface reactivity (*48*), which leads to extraordinary volumetric capacitances (>700 F/cm^3; Fig. 3C) (*49*). Yet in both MXenes and MOFs, the high surface area accelerates degradation, and the narrow pore throats limit mass transfer (*50*). High surface area may promote undesired reactivity in energy devices, such as corrosion, passivation, crack initiation, and energy dissipation, which stresses the need for fine-tuning surface area versus other energy transfer vectors.

Mass transfer in fuel cells and electrolyzers

High power is a key objective in electrochemical power sources, such as fuel cells, electrolyzers, and batteries. This calls for optimizing the pore structure for fast mass transport of reactants and products to and from a high volumetric surface area while maintaining electronic and heat conduction through the solid and across the solid-solution interface. Most mass transfer occurs by bulk and Knudsen diffusion inside pores and cracks, driven by the concentration gradients caused by reactant consumption and product formation. Convection plays a role when pumping (e.g., redox flow batteries), and migration occurs in insulating solids (e.g., lithium-ion batteries).

In proton exchange membrane fuel cells, multisized pores distribute the flow of reactants and products. O_2 gas arriving at the cathode passes through porosity-graded layers (Fig. 3D) (*51*). The flow is distributed by machined channels in a bipolar plate (width, 0.6 to 1 mm) into a gas diffusion layer (GDL), typically composed of fibrous carbon paper [d_{pore}, 10 to 50 μm (*52*)], and then to finer microporous layers (MPLs), whose textural porosity arises from the packing of carbon particles and polymeric binders (d_{pore}, 20 to 500 nm). When the O_2 reaches a catalyst layer (CL) of platinum nanoparticles on microporous carbon [d_{pore}, 1 to 50 nm (*53*)], it is reduced ($O_2 + 4H^+ + 4e^- \rightarrow 2H_2O$). Porosity is crucial for exposing the catalyst nanoparticles to reactants (*54*), protecting them from poisoning by the ionomer (*55*), and evacuating the produced water backward through the same sequence of layers (lest it condenses and blocks access to incoming O_2). Designed pores can help retain water at the CL-GDL interface, abating membrane dehydration. Natural drainage channels, such as textural porosity and random cracks, may be sufficient or detrimental, depending on operating conditions (Fig. 3E) (*56*). "Hierarchical" (size-graded) porosity can be introduced to different layers for managing water transport by electrospinning differently spaced fibers (*57*), deliberate cracking (*58*), laser grooving (*59*, *60*), or other nanofabrication methods (*61*, *62*). A thickness-tuned MPL can switch the oxygen flow pattern in a polymer electrolyte membrane water electrolyzer from capillary to viscous fingering, improving the electrolyte-electrode contact and exposing active sites (Fig. 3E).

For mass transport, the pores must be large enough for the diffusing species (O_2, H_2O, or ions), open to the external world, and interconnected; however, the introduction of porosity must minimize both the loss of volumetric surface area $\left(\mathrm{VSA} \propto 1/d_{pore}^2\right)$ and the loss of bulk matter, which conducts charge, heat, and absorbing light $\left(\varepsilon \propto 1/d_{pore}^3\right)$. These requirements define the "void triangle," where mass transfer is possible (Fig. 4A). In electrochemical devices and in oil-rich rocks, the pore dimensions determine the balance between the surface tension and capillary forces wicking the water out (*56*) or breaking water into pores (*11*). Because the water pressure for entering a pore scales inversely with d_{throat} (per Young-Laplace above), the layer hydrophobicity can be enhanced (and O_2 delivery improved) by engineering the pore curvature (*11*). Paradoxically, limiting mass transfer may enhance the yield of multistep electrochemical reactions—for example, O_2 (*63*) or CO_2 (*64*) reduction intermediates are retained in mesopores (30 to 50 nm) to react fully.

The porosity of electrodes that consume or produce gases, such as water electrolyzers (*65*), electrosynthetic reactors (*66*), or direct liquid fuel cells (*67*), must be tuned to balance the multiphase fluid transfer. High-temperature devices, such as phosphoric acid fuel cells, face specific challenges on shutoff, when condensation leads to pore blocking (*68*). Beyond pore size, fast gas diffusion in a porous electrode (D_{eff}) calls for minimal tortuosity (τ, ratio between the shortest possible pathway and the Euclidean distance between points), $D_{eff} = D \cdot \varepsilon / \tau$ (where D is the species diffusivity). Tortuosity is estimated as $\tau = \varepsilon^{-0.5}$ by the Bruggeman model for mass transport around randomly distributed spheres; yet, this is a poor approximation for fibrous and ionomer-embedded GDLs (*69*). The water distribution in electrodes is visualized and modeled on multiple length-scales by microscopic, spectroscopic, and tomographic techniques (*65*, *70*), some of which borrowed directly from geoscience. X-ray tomographic microscopy of commercial MPL-GDL composites reveal the spatial distribution of heterogeneities,

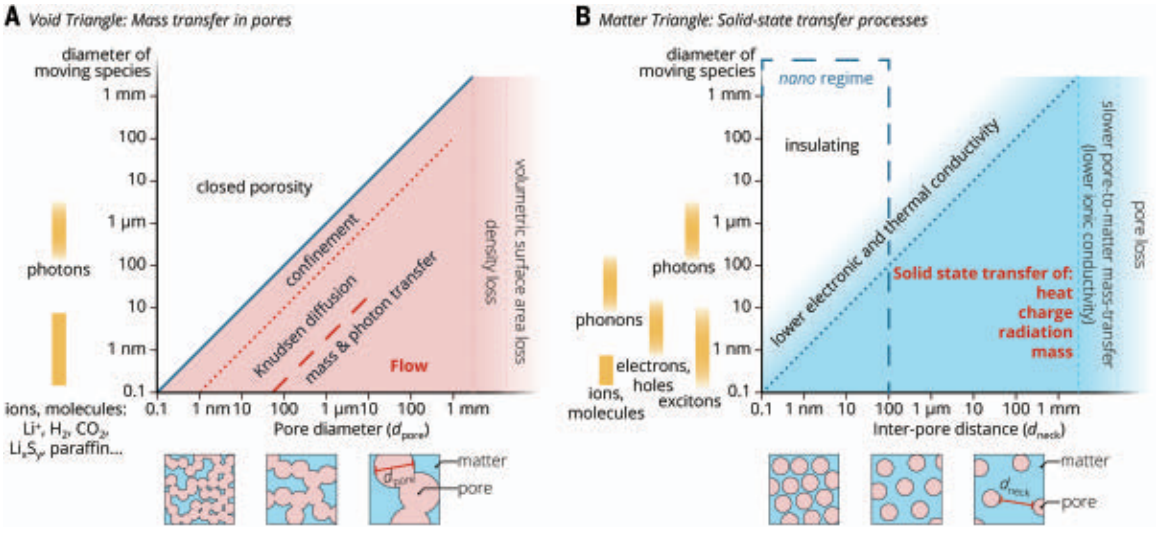


Fig. 4. Mapping transfer processes in porous structures. (**A**) The void triangle: A map of mass transfer regimes and conditions. The diagonal limit is set by the size of the moving species (unsolvated ion, molecule, or photon; $d_{photon} \sim \frac{5}{\pi}\lambda$) relative to the pore diameter. The vertical limit marks the disappearance of matter. Mass transfer regimes include flow (by bulk diffusion, convection, or migration), Knudsen (surface) diffusion for pores smaller than the MFP, and confinement. These processes occur in most applications discussed in this Review, from geology to electro-, thermo-, and photochemistry. (**B**) The matter triangle: A map of solid-state heat, charge, radiation, and mass transfer regimes and conditions. The diagonal limit is set by the size of the moving species relative to interpore distance (d_{neck}), but it is less strict than in (A) because some species (e.g., photons, phonons, excitons, and electrons) can diffract around the pores. The vertical limit marks the disappearance of pores. In the nano-regime, the matter dimensions give rise to energy quantization. Both triangles assume a random distribution of spherical pores (connected and disconnected, respectively); the maps would look different under other assumptions.

such as cracks, holes, and variations in thickness and density (Fig. 3F) (*56*), leading to preferential pathways for reactant flow (*71*). Visualization methods can be coupled with measurements of wetting angle distributions (*70*) or flow pathways (*72*) and analyzed computationally (*73*, *74*), a strategy inspired by the geosciences.

Mass and charge transfer inside battery electrodes

In batteries, deployed from portable to grid energy storage, pores throughout the electrodes can be filled with electrolyte to facilitate ion transport across an enhanced electrode surface. Batteries have higher energy density than supercapacitors (because ions fill the solid volume too) but lower power density because of slower mass transfer (Fig. 3G) (*75*). Battery porosity is typically textural ($\varepsilon = 0.3$ to 0.4), depending on the sizes and shapes of the packed particles. Templating the pores allows tuning both the sizes of pores and their wall thickness: Thinner walls shorten ion diffusion lengths in the solid and improve high-rate capabilities (*76*). To achieve "extreme fast charging," electrode particles are downsized to shorten intraparticle diffusion lengths of intercalating ions (e.g., Li^+); yet, this increases the tortuosity, slowing diffusion through interparticle voids (*77*). Textural porosity can be tweaked to improve both electronic and ionic conductivity of lithium titanate—a safe alternative to graphite in Li-ion batteries—by nanometric sculpting of the particles (*78*), and a similar approach is taken in fuel cell electrodes. Alternatively, porous single crystals were designed to reduce energy loss at particle boundaries, providing thin walls for fast Li^+ diffusion and a porous network (formed by removing ~5-nm octahedra) to speed electrolyte diffusion (Fig. 3H) (*79*). Electronic conductivity can also be enhanced by tuning the ligand structure of MOF electrodes (*80*).

Well-ordered porosity (e.g., in inverse opals and carbon frameworks) helps homogenize the electric field in metal-air batteries, enhancing mass transfer and mitigating the growth of dendrites that short the device (Fig. 3I) (*81*). However, such high porosity lowers the volumetric surface area and promotes unwanted formation of electrode-blocking interphases. In contrast to such devices, where the porosity of the metal electrode changes during operation, metal sulfur batteries store the active sulfur inside the fixed and tunable pores of a carbon matrix. In this case, pore throats must be tuned to minimize leaching of lithiated polysulfide chains (Li_XS_y) (*76*, *82*), to protect the sulfur from pulverization during cyclic volume changes (*83*) and to enable sufficient electrical conductivity to the carbon (*84*). Overall, d_{pore} determines the amount of stored active material (energy density), and d_{throat} determines the mass transfer rate (power density) as well as loading and leaching efficiency (which contribute to energy density). Active materials are also encapsulated in porous heat storage reactors, which calls for an exchange of matrix materials and porosity engineering methods between the fields.

The "matter triangle" (Fig. 4B), describing the limits and conditions for heat, charge, radiation, and mass transfer in the solid matrix, is a conceptual mirror of the void triangle (Fig. 4A), which shows mass and radiation transfer in the pores. As seen in batteries and thermochemical applications (see "Thermal heat storage" section), charge and heat carriers move more easily at larger interpore distances (d_{neck}) but at the expense of slower diffusion between the pores and the solid [$t \propto D / d^2_{neck}$, or less if defects block diffusion channels (*85*)]. In both the matter and void triangles, the porous nature of the material is destroyed by an excess of matter or voids, respectively (the right vertical). All transfer processes become limited when carriers are larger than throats (d_{neck} or d_{pore}, respectively; the diagonal), although solid-state energy carriers, such as phonons and electrons, can diffract around pores. Below $d_{neck} < 100$ nm, quantum confinement in the solid creates a distinct nano-regime, somewhat mirroring intermediate confinement in supercapacitors (*40*), electrocatalysts (*86*), or subsurface oil (*29*). In this regime, higher strains created by nano-pores improve the miscibility of Li^+ in battery materials (*85*) while enhancing resistance to cracking upon volume expansion (*27*).

Heat and charge transfer in thermoelectric generators

In thermoelectric generators, two dissimilar materials (e.g., a *p*-type and an *n*-type semiconductor) produce a voltage by the Seebeck effect when subjected to a temperature difference (Fig. 5A). The dimensionless figure of merit for thermoelectric conversion is $ZT = \frac{S^2 T \sigma}{\kappa_{phon,sol} + \kappa_{e^-,sol} + \kappa_{MT,por} + \kappa_{cond,por} + \kappa_{RT,por}}$, where S is the Seebeck coefficient, σ is electrical conductivity, and κ is thermal conductivity caried in the solid (by phonons, $\kappa_{phon,sol}$, and electrons, $\kappa_{e^-,sol}$) or in the pores (by convection, $\kappa_{MT,por}$, and conduction, $\kappa_{cond,por}$, by pore-filling fluids, or by radiation in the pores, $\kappa_{RT,por}$). This equation reveals a built-in contradiction of designing thermoelectric solids: Low thermal conductivity is needed for improving the heat gradient, whereas high electrical conductivity is required for increasing power; unfortunately, the same charge carrier (electrons) can conduct both heat and electricity (*87*).

Porosity emerges as a detour around this conundrum (*88*). Because heat-conducting electrons and phonons (lattice vibrations) differ in mean free path (MFP) (~1 to 10 and ~10 to 100 nm, respectively), pores can be tuned in size to scatter phonons while letting electrons through (*89*). Fine-tuning the pore fraction in Si nanowires lowered κ without compromising σ too much, overall boosting ZT (Fig. 5B) (*90*). The carrier depletion that porosity introduces can be modeled by $x_{por} = x_{bulk}\frac{1-\varepsilon}{1+\varepsilon}$, where x is either κ or σ (*91*); this is also a useful first approximation for mass transport in electrochemical devices. More complex models, taking into account pore size, shape, orientation, and spatial distribution, predict a drop in phononic heat conductivity when $\varepsilon > 0.75$ (*92*). The heat transport in a porous matrix also depends on phonon size $\left[\kappa_{solid} \propto MFP_{phonon}(1-\varepsilon)\right]$ and on convection $\left(\kappa_{MT} \propto \frac{1}{d^n_{pore}}\right)$ and conduction $\left(\kappa_{fluid} \propto \frac{d_{pore}}{MFP_{fluid}}\right)$ by any pore-trapped fluids. Thus, introducing closed pores (*93*) or pores smaller than the MFP of air (~70 nm), helps reduce κ.

Nanoporous thermoelectric materials, from Bi_2Te_3 alloys to lead chalcogenides, display lower κ and higher ZT values (*94*, *95*) because smaller pores mean a shorter MFP_{phonon} and a sharper decline in κ (*96*). Pores oriented perpendicularly to the heat transport direction can block more phonons, as do interconnected pores (at the expense of stronger electrical isolation) (*97*). Periodic porosity blocks phonons while allowing the electrons to travel in a straight line (*96*). Although low-density MOFs are ideal for thermal insulation, their thermoelectric performance still lags behind (*98*). A Zr-based MOF with both tetrahedral and octahedral pores (0.8 and 1.1 nm, respectively) was used to host pyrrole monomers selectively in the larger pores, yielding, upon polymerization, a network of conductive polypyrrole wires through the material (Fig. 5C). This increased σ a million-fold while mostly retaining the low κ thanks to the porosity of the smaller voids (*99*).

Thermal heat storage

Latent heat storage uses phase change materials (PCMs) hosted in porous matrices (*100*), whereas thermochemistry converts heat by reversible endothermic reactions of porous or pore-encapsulated materials (*101*). The heat stored is $Q = X \cdot n \cdot \Delta H_r$, where X is the reaction conversion (1 for phase change, less for thermochemistry), n is the number of moles, and ΔH_r is the molar reaction enthalpy.

PCMs are best suited for waste heat and concentrated solar storage, typically storing 20 to 500 kJ/kg at temperatures of 50° to 600°C (Fig. 5D) (*102*). To improve the heat conductivity through a packed bed reactor of thermally insulating PCM particles, the material (such as paraffin wax) can be stored inside porous aerogels (*103*), sponges (*104*), or microporous polymers (*105*). As in metal sulfur batteries, the infiltration of a PCM into the porous matrix requires open porosity, leading to leaching during operation (*106*). Alternatively, cosynthesis of nanoparticles (e.g., Sn or Pb) in a mesoporous carbon is achieved by pyrolysis, yielding well-dispersed and stably anchored PCMs and fast heat transfer across the large solid-solid interfacial area (*107*). The mesopores confine

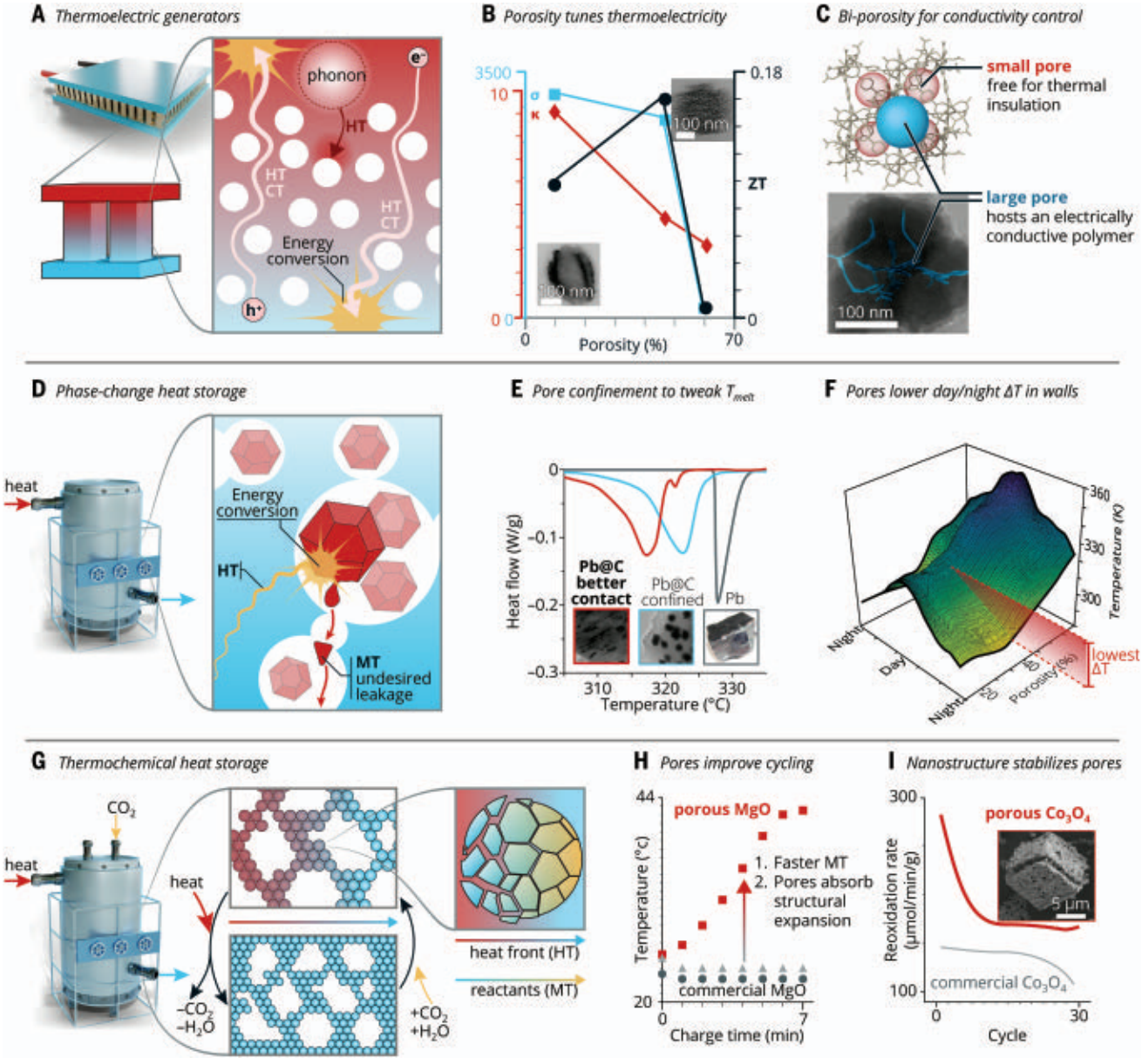


Fig. 5. Thermal energy conversion by porous materials in thermoelectric generators, latent heat storage PCMs, and thermochemical reactors. (**A** to **C**) Thermoelectric generators. (**D** to **F**) Latent heat storage PCMs. (**G** to **I**) Thermochemical reactors. The operational principles point to the transport phenomena occurring in each application, including mass transfer (MT), charge transfer (CT), and heat transfer (HT). (B) High-temperature thermoelectricity in Si nanowires is enhanced by tuning porosity to lower thermal conductivity without degrading electrical conductivity (*90*). (C) A bicontinuous MOF enables ultralow thermal conductivity, and partial filling of pores with a conductive polypyrrole polymer enhances the electronic conductivity a million-fold (*99*). (E) Lead nanoparticles PCM form inside a mesoporous carbon without postsynthetic impregnation, and their melting point can be tweaked by varying the nanoparticle size (*107*). (F) Multilevel simulations of heat transfer inside a hierarchically porous solar wall are crucial for optimizing this passive house heating technology (*110*). (H) Cycling in the $Mg(OH)_2$/MgO couple can be enhanced and stabilized if the material is prepared as high–surface area flower-shaped particles (*121*). (I) Porosity designed by partial sintering of nanostructured materials increases and steadies the thermochemical reoxidation rates (*123*).

the volume expansion, decreasing mechanical stresses and interfering with their crystallization (lowering the melting temperature). Such supercooling can tweak the operating temperature to the temperature of a home or a factory (*108*) at the expense of some latent heat (*109*). In the Pb/C system, suited for higher temperatures, the melting point could be tweaked by tuning the metal particle sizes, thus varying the metal surface:volume ratio and the metal-carbon interface area (Fig. 5E) (*107*). A porosity-within-porosity concept was implemented in solar walls, absorbing sunlight at daytime and releasing heat at night, where 10-mm carbon fibers were impregnated with 10-μm polymer-encapsulated paraffin wax (*110*). Heat and mass transfer were optimized numerically to minimize the day-to-night difference ΔT (Fig. 5F). Larger throat diameters (>100 nm) are desired for better heat conductivity, giving porous carbons an advantage over MOFs as thermochemical hosts. Both fields of PCMs and electrochemistry use carbons, offering better pore tuneability than brittle and hard inorganic oxides or carbonates. Moreover, pore-collapsing sintering of most thermochemical inorganic oxides begins in the 400° to 800°C range, whereas the micropores of a thermoset carbon only start shrinking around 800° to 1200°C (*111*).

In thermochemical cycling, an inorganic material may store heat by partial decomposition to produce a gas (e.g., H_2O, CO_2, or O_2) and release the heat by reacting with the gas again (Fig. 5G). High energy

densities (<0.9 MJ/kg) (*112*) are achieved at high operating temperatures (<1500°C) (*102*). The active material is either porous itself or encapsulated in a porous matrix (*113*) to enable fast mass transfer of reactive gases to a high surface area (*114*). Many thermochemical loops use abundant and naturally porous alkaline earth minerals (*115*). The cycle $CaCO_3(s) \rightleftharpoons CaO(s)+CO_2(g)$ ($\Delta H° = 178$ kJ/mol), where both calcination (energy storage) and carbonation (energy release) proceed best at high temperatures (700° to 900°C), is ideal for storing solar thermal radiation (*116*). Understanding the reaction fronts and pore dynamics in thermochemical materials is both important and challenging because all transfer processes are coupled (Fig. 4B) (*117*). Ca-based materials lose porosity rapidly during operation by pore sintering (*118*) or by particle attrition due to cyclic mechanical stresses (*119*). This leads to complex reaction fronts within the porous reactive media (*116*). Powders can be stabilized against attrition by shaping into porous forms (*120*), but typically up to 50% of the volumetric energy density is lost within 20 to 30 cycles (*116*). In both heat-driven and electrochemical applications, a desire for high reactive surface area must be balanced against chemo-mechanical degradation and tortuosity. Solutions such as hierarchical porosity of MgO flowers prevent mechanical stresses during the $Mg(OH)_2(s) \rightleftharpoons MgO(s)+H_2O(g)$ cycle ($\Delta H° = 81$ kJ/mol, 400°C), thanks to expansion buffering by the interwall gaps, thinner MgO walls, and improved conductivity (Fig. 5H) (*121*).

Cobalt oxide thermochemistry uses O_2, obviating the need for separate gas storage: $Co_3O_4(s) \rightleftharpoons 3CoO(s)+{}^1/_2O_2(g)$ ($\Delta H° = 196$ kJ/mol), with high energy density (844 kJ/kg), fast reaction kinetics, and high conversion (*122*). Carefully designed porosity enhances reactivity and mass transfer (*123*) and can even stabilize against pore collapse (*123*). For example, cube-shaped Co_3O_4 particles undergo partial sintering, forming a stable porous structure that keeps breathing O_2, whereas commercial Co_3O_4 powder collapses (Fig. 5I). The porosity of cobalt oxides can be enhanced by doping (*124*) or by coating on a foam-shaped cordierite mineral, which operates for 100 cycles (*125*), whereas Co_3O_4 pellets crack within 10 cycles.

Radiation transfer in solar cells and reactors

Solar radiation can be harvested by photovoltaic panels or thermochemical converters. Conversion of photons to voltage or heat occurs in the solid phase, but porosity governs light harvesting and scattering (including diffraction, reflection, and refraction). Light harvesting in solar cells may be improved by reflectors or absorbers made of porous silicon, which enhance light scattering (*126*). Metamaterials based on pore arrays have been shown to confine photons, improving harvesting and scattering behavior. Periodically porous light-trapping layers, such as inverse opal TiO_2, reduce the group velocity of photons at stop band edges (*127*), suppressing electron-hole recombination and boosting light conversion efficiency compared with a disordered macroporous film (*128*). The control of radiation transfer by porous structures is also exploited in photoelectrocatalysis (*129*). Simulations show that the pore sizes (*130*) and total porosity (*131*) determine refractive indices as well as the light absorption peaks (value, intensity, and shape), creating a photonic crystal that can capture and guide light to the device (*132*).

In dye-sensitized solar cells, mesopores formed between packed TiO_2 nanocrystals in the electron transport layers allow the passage of electrolyte redox couples (*133*). In perovskite-based solar cells, porous transport layers host the perovskite crystals [affecting their size, crystallinity, and thus optical properties (*134*)], and their roughness improves contact to the perovskite layer, shortening the charge conduction pathways (*135*). Small pores improve the charge extraction from the perovskite layer by tightening the contact to the transport layers and can also scatter incoming light, lengthening its path in the material and boosting overall light harvesting efficiency (*136*). MOFs also serve as the electron- or hole-transport layer or are embedded in either. To address the pressing issue of stability, Pb^{2+} ions leaching from the perovskite layer were caught in thiol-lined pores of a Zr-MOF embedded in the electron transport layer (Fig. 6, B and C) (*137*), whereas pyridyl-lined pores in an In-MOF captured leaching Li^+ (*138*). When the perovskite material is hosted in MOF pores, charge transfer through the scaffold is enhanced, and the growth of

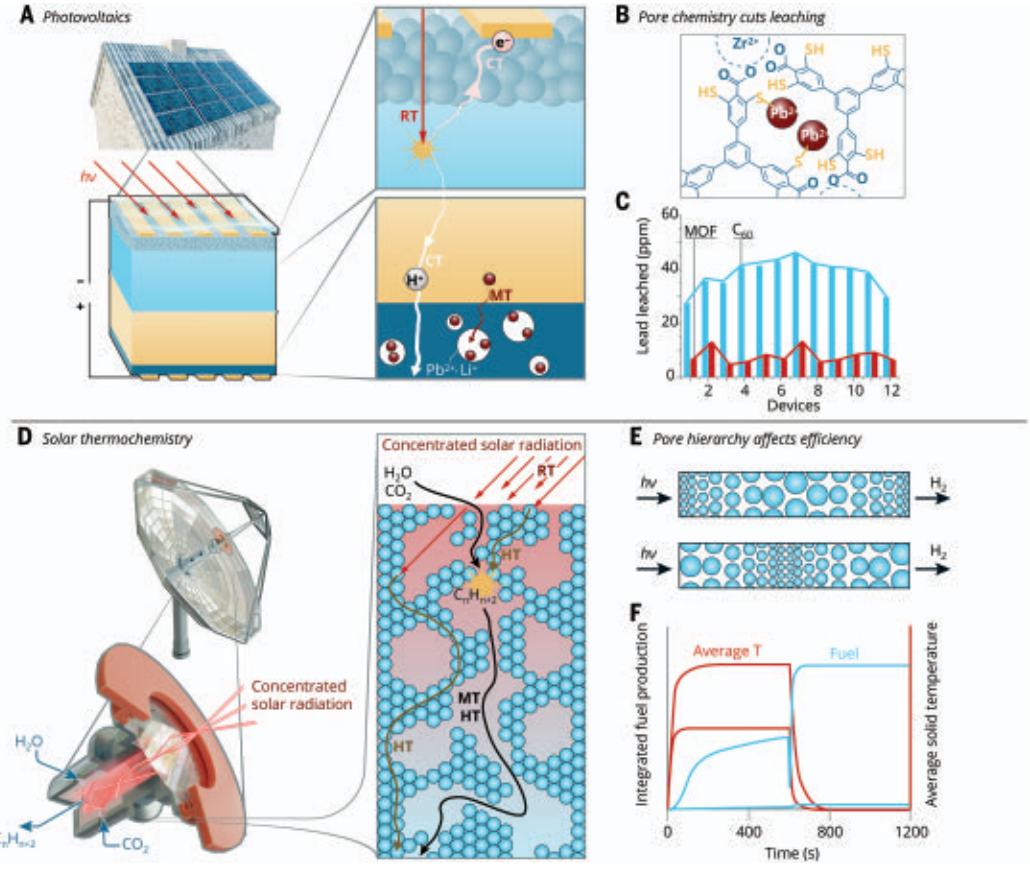


Fig. 6. Solar energy conversion by porous materials in perovskite photovoltaics, and solar thermochemical reactors. (**A** to **C**) Perovskite photovoltaics. (**D** to **F**) Solar thermochemical reactors. The operational principles point to the transport phenomena occurring in each application, including mass transfer (MT), charge transfer (CT), heat transfer (HT), and radiation transfer (RT). (B) The thiol-lined pores of a Zr-MOF embedded in the electron transfer layer can catch Pb^{2+} ions leaching from the perovskite layer (*137*), leading to (C) better Pb retention. (E) Optimal pore gradients were identified by experiment and simultaneous optimization of light penetration and heat and mass transfer, leading to (F) improved solar fuel production (*143*).

the perovskite crystals is moderated, boosting the cell efficiency of the solar cell (*139*). Alternatively, a MPL of Zn-MOF can store and release reactive salts gradually during perovskite growth, compensating for vacancy defects (*140*). The storage abilities of MOFs could be exploited for storing active materials in batteries (*141*) and heat storage devices (*142*).

Solar thermal converters use porous ceramics as receivers to store solar heat radiation in chemical bonds at high temperatures (>1000°C; Fig. 6D). For example, macroporous ceria (CeO_2) can be endothermically reduced at ~1600°C and then catalytically reoxidized by incoming H_2O or CO_2 at ~1100°C. Theory predicts a thermochemical solar-to-fuel efficiency of 35%, but experimental reactors rarely reach 10% (*143*). At these temperatures, radiation through the pores dominates heat transfer, improving at higher porosity ($\varepsilon < 0.9$) and boosting fuel yield at the expense of active material loading; the optimum for overall productivity is $\varepsilon = 0.6$ to 0.75 (*143*). High incident radiation requires large pores and high porosity to homogenize the temperature. Grading the pores by size (Fig. 6E) may enhance photon influx and reduce radiation loss at the back side (Fig. 6F) (*144*), although reactor design may prove porosity grading unnecessary (*145*) or even detrimental (*146*).

Approaches are being developed to understand the simultaneous transfer of radiation, heat, and mass at such high illumination intensities. Beyond treating a porous material as a lower-density nonporous one, the radiation transport in pores can be analyzed by geometrical optics, valid when pores are large enough relative to the photon wavelength, roughly $\pi d_{\text{pore}} > 5\lambda$ (*147*). Although such models neglect scattering, they yield the effective transfer properties by semiempirical combinations of real porous geometries (from 3D tomographic imaging) with numerical simulations (by Monte Carlo or ray tracing methods) (*148*). The fullest, and computationally heaviest, descriptions of radiation transport come from pore scale models that solve the Maxwell equations explicitly. This approach captures the effects of light trapping in the regime when $\lambda \sim d_{\text{pore}}$ (*149*). Such modeling methods can also help solve local hotspots in fuel cell electrodes caused by inhomogeneous reactivity or current distribution, which degrade the material and complicate device operation (*150*).

Radiation-induced porosity occurs in nuclear reactors, where incoming neutron irradiation produces gaseous nuclear fission products (Xe and Kr) in the nuclear fuel pellet ($^{235}UO_2$) (*151*). To release these bubbles, lest they block heat removal and further neutron propagation, the pellet must have interconnected porosity. Thus, the choice of fuel composition and crystallinity, along with radiation intensity, determines the evolving porosity and, in turn, the radiation, heat, and mass transfer in the material (*152*). This situation is somewhat analogous to the spontaneous (and often undesired) porosity that forms upon metal deposition in batteries (*153*).

The future of porosity research in energy technologies

The ultimate goal in porosity research is to predict an optimal pore structure for each energy application. This calls for multiscale computational models (*154*) and mathematical tools (*155*) to describe and predict coupled energy transfer processes. New ideas for synthetic regulation of porosity, such as controlled twisting of packed polymer chains (*156*) or zeolite exfoliation (*157*), require multiscale characterization techniques, including extending computerized tomography (*158*) or nuclear magnetic resonance cryoporometry (*159*) to the nanoscale. Such cross-fertilization of imaging methods is already affecting thermal catalysis (*160*) and electrocatalysis (*70*) and could benefit from more biomimetic inspirations (*161*). Porous materials could be shared more between different applications—e.g., carbon electrocatalysts as PCM matrices or oxide photoelectrodes as thermochemical reactants. Other energy transfer processes can be considered, such as magnetocaloric cooling in oxide foams (*162*), piezoelectricity and triboelectricity in porous energy harvesters (*163*), or directional mass transfer in water-purifying membranes (*164*). Porous materials can find new energy applications, such as mechanical energy absorption in titanium bone implants with gradient porosity (*165*) or integrated circuits with reduced power consumption using porous low-dielectric polymers (*166*). Ultimately, when porous structures can be precisely designed, fully described, and modeled on all length- and timescales, their potential for energy devices will be truly realized.

REFERENCES AND NOTES

1. D. Eisenberg, Ed., *Handbook of Porous Materials* (World Scientific, 2021).
2. R. M. M. Schneerson, *Likkutei Sichos (Collected Talks)* (Kehot, 2016).
3. I. Ligaarden, M. Krotkiewski, K. A. Lie, M. Pal, D. Schmid, "On the Stokes-Brinkman Equations for Modeling Flow in Carbonate Reservoirs" (European Association of Geoscientists & Engineers, 2010); https://www.earthdoc.org/content/papers/10.3997/2214-4609.20144924.
4. R. D. Wyckoff, H. G. Botset, The Flow of Gas-Liquid Mixtures Through Unconsolidated Sands. *Physics* **7**, 325–345 (1936). doi: 10.1063/1.1745402
5. D. G. Avraam, A. C. Payatakes, Flow regimes and relative permeabilities during steady-state two-phase flow in porous media. *J. Fluid Mech.* **293**, 207–236 (1995). doi: 10.1017/S0022112095001698
6. R. T. Armstrong *et al.*, Flow Regimes During Immiscible Displacement. *Petrophysics* **58**, 10–18 (2017).
7. C. Spurin *et al.*, The development of intermittent multiphase fluid flow pathways through a porous rock. *Adv. Water Resour.* **150**, 103868 (2021). doi: 10.1016/j.advwatres.2021.103868
8. F. Hoogland, P. Lehmann, R. Mokso, D. Or, Drainage mechanisms in porous media: From piston-like invasion to formation of corner flow networks. *Water Resour. Res.* **52**, 8413–8436 (2016). doi: 10.1002/2016WR019299
9. Z. Sun, J. C. Santamarina, Haines jumps: Pore scale mechanisms. *Phys. Rev. E* **100**, 023115 (2019). doi: 10.1103/PhysRevE.100.023115; pmid: 31574718
10. M. Rücker *et al.*, From connected pathway flow to ganglion dynamics. *Geophys. Res. Lett.* **42**, 3888–3894 (2015). doi: 10.1002/2015GL064007
11. C. Simon, J. Endres, B. Nefzger-Loders, F. Wilhelm, H. A. Gasteiger, Interaction of Pore Size and Hydrophobicity/Hydrophilicity for Improved Oxygen and Water Transport through Microporous Layers. *J. Electrochem. Soc.* **166**, F1022–F1035 (2019). doi: 10.1149/2.1111913jes
12. J. Feng *et al.*, An end-to-end three-dimensional reconstruction framework of porous media from a single two-dimensional image based on deep learning. *Comput. Methods Appl. Mech. Eng.* **368**, 113043 (2020). doi: 10.1016/j.cma.2020.113043
13. Q. Xiong, T. G. Baychev, A. P. Jivkov, Review of pore network modelling of porous media: Experimental characterisations, network constructions and applications to reactive transport. *J. Contam. Hydrol.* **192**, 101–117 (2016). doi: 10.1016/j.jconhyd.2016.07.002; pmid: 27442725
14. I. Fatt, The Network Model of Porous Media. *Trans. AIME* **207**, 144–181 (1956).
15. C. Sun *et al.*, Characterization of wetting using topological principles. *J. Colloid Interface Sci.* **578**, 106–115 (2020). doi: 10.1016/j.jcis.2020.05.076; pmid: 32521350
16. R. T. Armstrong *et al.*, Porous Media Characterization Using Minkowski Functionals: Theories, Applications and Future Directions. *Transp. Porous Media* **130**, 305–335 (2019). doi: 10.1007/s11242-018-1201-4
17. F. Al-Zubaidi *et al.*, Effective permeability of an immiscible fluid in porous media determined from its geometric state. *Phys. Rev. Fluids* **8**, 064004 (2023). doi: 10.1103/PhysRevFluids.8.064004
18. A. Hansen, Linearity of the Co-moving Velocity. *Transp. Porous Media* **151**, 2477–2489 (2024). doi: 10.1007/s11242-024-02121-9
19. D. Bedeaux, S. Kjelstrup, Fluctuation-Dissipation Theorems for Multiphase Flow in Porous Media. *Entropy* **24**, 46 (2021). doi: 10.3390/e24010046; pmid: 35052072
20. D. T. Birdsell, H. Rajaram, D. Dempsey, H. S. Viswanathan, Hydraulic fracturing fluid migration in the subsurface: A review and expanded modeling results. *Water Resour. Res.* **51**, 7159–7188 (2015). doi: 10.1002/2015WR017810
21. M. S. Liew, K. U. Danyaro, N. A. W. A. Zawawi, A Comprehensive Guide to Different Fracturing Technologies: A Review. *Energies* **13**, 3326 (2020). doi: 10.3390/en13133326
22. J. Jello, T. Baser, Utilization of existing hydrocarbon wells for geothermal system development: A review. *Appl. Energy* **348**, 121456 (2023). doi: 10.1016/j.apenergy.2023.121456
23. G. Lu, G. Q. Lu, Z. M. Xiao, Mechanical Properties of Porous Materials. *J. Porous Mater.* **6**, 359–368 (1999). doi: 10.1023/A:1009669730778
24. K. Narita, M. A. Citrin, H. Yang, X. Xia, J. R. Greer, 3D Architected Carbon Electrodes for Energy Storage. *Adv. Energy Mater.* **11**, 2002637 (2021). doi: 10.1002/aenm.202002637
25. D. A. Wood, B. Hazra, Characterization of organic-rich shales for petroleum exploration and exploitation: A review-Part 3: Applied geomechanics, petrophysics and reservoir modeling. *J. Earth Sci.* **28**, 779–803 (2017). doi: 10.1007/s12583-017-0734-8
26. L. A. B. da Cunda, G. J. Creus, in *Plasticity of Pressure-Sensitive Materials*, H. Altenbach, A. Öchsner, Eds. (Springer, 2014), pp. 349–376.
27. X. H. Liu *et al.*, Size-dependent fracture of silicon nanoparticles during lithiation. *ACS Nano* **6**, 1522–1531 (2012). doi: 10.1021/nn204476h; pmid: 22217200

28. S. G. Rudisill, Z. Wang, A. Stein, Maintaining the structure of templated porous materials for reactive and high-temperature applications. *Langmuir* **28**, 7310–7324 (2012). doi: 10.1021/la300517g; pmid: 22409622
29. J. E. Santos *et al.*, Modeling Nanoconfinement Effects Using Active Learning. *J. Phys. Chem. C* **124**, 22200–22211 (2020). doi: 10.1021/acs.jpcc.0c07427
30. L. Brochard, M. Vandamme, R. J.-M. Pellenq, Poromechanics of microporous media. *J. Mech. Phys. Solids* **60**, 606–622 (2012). doi: 10.1016/j.jmps.2012.01.001
31. J. M. Baptista, J. S. Sagu, U. W. Kg, K. Lobato, State-of-the-art materials for high power and high energy supercapacitors: Performance metrics and obstacles for the transition from lab to industrial scale – A critical approach. *Chem. Eng. J.* **374**, 1153–1179 (2019). doi: 10.1016/j.cej.2019.05.207
32. D. P. Chatterjee, A. K. Nandi, A review on the recent advances in hybrid supercapacitors. *J. Mater. Chem. A* **9**, 15880–15918 (2021). doi: 10.1039/D1TA02505H
33. E. Frackowiak, Q. Abbas, F. Béguin, Carbon/carbon supercapacitors. *J. Energy Chem.* **22**, 226–240 (2013). doi: 10.1016/S2095-4956(13)60028-5
34. Z. Li *et al.*, Tuning the interlayer spacing of graphene laminate films for efficient pore utilization towards compact capacitive energy storage. *Nat. Energy* **5**, 160–168 (2020). doi: 10.1038/s41560-020-0560-6
35. J. Chmiola *et al.*, Anomalous increase in carbon capacitance at pore sizes less than 1 nanometer. *Science* **313**, 1760–1763 (2006). doi: 10.1126/science.1132195; pmid: 16917025
36. E. Raymundo-Piñero, K. Kierzek, J. Machnikowski, F. Béguin, Relationship between the nanoporous texture of activated carbons and their capacitance properties in different electrolytes. *Carbon* **44**, 2498–2507 (2006). doi: 10.1016/j.carbon.2006.05.022
37. K. Ge, H. Shao, Z. Lin, P.-L. Taberna, P. Simon, Advanced characterization of confined electrochemical interfaces in electrochemical capacitors. *Nat. Nanotechnol.* **20**, 196–208 (2025). doi: 10.1038/s41565-024-01821-z; pmid: 39639178
38. H. Ma *et al.*, Maximization of Spatial Charge Density: An Approach to Ultrahigh Energy Density of Capacitive Charge Storage. *Angew. Chem. Int. Ed.* **59**, 14541–14549 (2020). doi: 10.1002/anie.202005270; pmid: 32506611
39. J. Huang *et al.*, Giant Redox Entropy in the Intercalation vs Surface Chemistry of Nanocrystal Frameworks with Confined Pores. *J. Am. Chem. Soc.* **145**, 6257–6269 (2023). doi: 10.1021/jacs.2c12846; pmid: 36893341
40. S. Fleischmann *et al.*, Continuous transition from double-layer to Faradaic charge storage in confined electrolytes. *Nat. Energy* **7**, 222–228 (2022). doi: 10.1038/s41560-022-00993-z
41. K. V. G. Raghavendra *et al.*, An intuitive review of supercapacitors with recent progress and novel device applications. *J. Energy Storage* **31**, 101652 (2020). doi: 10.1016/j.est.2020.101652
42. J. Biemolt, I. M. Denekamp, T. K. Slot, G. Rothenberg, D. Eisenberg, Boosting the Supercapacitance of Nitrogen-Doped Carbon by Tuning Surface Functionalities. *ChemSusChem* **10**, 4018–4024 (2017). doi: 10.1002/cssc.201700902; pmid: 28586123
43. X. Zhu, Recent advances of transition metal oxides and chalcogenides in pseudo-capacitors and hybrid capacitors: A review of structures, synthetic strategies, and mechanism studies. *J. Energy Storage* **49**, 104148 (2022). doi: 10.1016/j.est.2022.104148
44. X. Li *et al.*, MXene chemistry, electrochemistry and energy storage applications. *Nat. Rev. Chem.* **6**, 389–404 (2022). doi: 10.1038/s41570-022-00384-8; pmid: 37117426
45. H. Shi *et al.*, 3D Flexible, Conductive, and Recyclable $Ti_3C_2T_x$ MXene-Melamine Foam for High-Areal-Capacity and Long-Lifetime Alkali-Metal Anode. *ACS Nano* **14**, 8678–8688 (2020). doi: 10.1021/acsnano.0c03042; pmid: 32530269
46. I. Ali, M. Yousaf, I. H. Sajid, M. W. Hakim, S. Rizwan, Reticulation of 1D/2D Mo_2TiC_2 MXene for excellent supercapacitor performance. *Mater. Today Chem.* **34**, 101766 (2023). doi: 10.1016/j.mtchem.2023.101766
47. Q. Tao *et al.*, Two-dimensional $Mo_{1.33}C$ MXene with divacancy ordering prepared from parent 3D laminate with in-plane chemical ordering. *Nat. Commun.* **8**, 14949 (2017). doi: 10.1038/ncomms14949; pmid: 28440271
48. D. Sheberla *et al.*, Conductive MOF electrodes for stable supercapacitors with high areal capacitance. *Nat. Mater.* **16**, 220–224 (2017). doi: 10.1038/nmat4766; pmid: 27723738
49. D. Feng *et al.*, Robust and conductive two-dimensional metal–organic frameworks with exceptionally high volumetric and areal capacitance. *Nat. Energy* **3**, 30–36 (2018). doi: 10.1038/s41560-017-0044-5
50. F. Bu *et al.*, Porous MXenes: Synthesis, structures, and applications. *Nano Today* **30**, 100803 (2020). doi: 10.1016/j.nantod.2019.100803
51. M. B. Sassin, Y. Garsany, R. W. Atkinson III, R. M. E. Hjelm, K. E. Swider-Lyons, Understanding the interplay between cathode catalyst layer porosity and thickness on transport limitations en route to high-performance PEMFCs. *Int. J. Hydrogen Energy* **44**, 16944–16955 (2019). doi: 10.1016/j.ijhydene.2019.04.194
52. F. C. Lee *et al.*, Alternative architectures and materials for PEMFC gas diffusion layers: A review and outlook. *Renew. Sustain. Energy Rev.* **166**, 112640 (2022). doi: 10.1016/j.rser.2022.112640
53. M. Uchida *et al.*, Effect of the state of distribution of supported Pt nanoparticles on effective Pt utilization in polymer electrolyte fuel cells. *Phys. Chem. Chem. Phys.* **15**, 11236–11247 (2013). doi: 10.1039/c3cp51801a; pmid: 23715296
54. G. S. Harzer, A. Orfanidi, H. El-Sayed, P. Madkikar, H. A. Gasteiger, Tailoring Catalyst Morphology towards High Performance for Low Pt Loaded PEMFC Cathodes. *J. Electrochem. Soc.* **165**, F770–F779 (2018). doi: 10.1149/2.0311810jes
55. V. Yarlagadda *et al.*, Boosting Fuel Cell Performance with Accessible Carbon Mesopores. *ACS Energy Lett.* **3**, 618–621 (2018). doi: 10.1021/acsenergylett.8b00186
56. Y.-C. Chen, C. Karageorgiou, J. Eller, T. J. Schmidt, F. N. Büchi, Determination of the porosity and its heterogeneity of fuel cell microporous layers by X-ray tomographic microscopy. *J. Power Sources* **539**, 231612 (2022). doi: 10.1016/j.jpowsour.2022.231612
57. M. Balakrishnan *et al.*, Designing Tailored Gas Diffusion Layers with Pore Size Gradients via Electrospinning for Polymer Electrolyte Membrane Fuel Cells. *ACS Appl. Energy Mater.* **3**, 2695–2707 (2020). doi: 10.1021/acsaem.9b02371
58. S. M. Kim *et al.*, High-performance Fuel Cell with Stretched Catalyst-Coated Membrane: One-step Formation of Cracked Electrode. *Sci. Rep.* **6**, 26503 (2016). doi: 10.1038/srep26503; pmid: 27210793
59. C. Csoklich, T. J. Schmidt, F. N. Büchi, High performance gas diffusion layers with added deterministic structures. *Energy Environ. Sci.* **15**, 1293–1306 (2022). doi: 10.1039/D1EE03246A
60. E. A. Chadwick, P. Shrestha, H. B. Parmar, A. Bazylak, V. P. Schulz, Biomimetic auxiliary channels enhance oxygen delivery and water removal in polymer electrolyte membrane fuel cells. *Appl. Energy* **389**, 125760 (2025). doi: 10.1016/j.apenergy.2025.125760
61. M. Breitwieser, M. Klingele, S. Vierrath, R. Zengerle, S. Thiele, Tailoring the Membrane-Electrode Interface in PEM Fuel Cells: A Review and Perspective on Novel Engineering Approaches. *Adv. Energy Mater.* **8**, 1701257 (2018). doi: 10.1002/aenm.201701257
62. C. Lee *et al.*, Grooved electrodes for high-power-density fuel cells. *Nat. Energy* **8**, 685–694 (2023). doi: 10.1038/s41560-023-01263-2
63. E. M. Farber, K. Ojha, T. Y. Burshtein, L. Hasson, D. Eisenberg, Understanding the self-templating of hierarchically porous carbon electrocatalysts using Group 2 coordination polymers. *Mater. Adv.* **1**, 20–33 (2020). doi: 10.1039/D0MA00084A
64. D. Hursán *et al.*, Morphological Attributes Govern Carbon Dioxide Reduction on N-Doped Carbon Electrodes. *Joule* **3**, 1719–1733 (2019). doi: 10.1016/j.joule.2019.05.007; pmid: 31417986
65. J. K. Lee *et al.*, Designing catalyst layer morphology for high-performance water electrolysis using synchrotron X-ray nanotomography. *Cell Rep. Phys. Sci.* **4**, 101232 (2023). doi: 10.1016/j.xcrp.2022.101232
66. J. C. Bui *et al.*, Continuum Modeling of Porous Electrodes for Electrochemical Synthesis. *Chem. Rev.* **122**, 11022–11084 (2022). doi: 10.1021/acs.chemrev.1c00901; pmid: 35507321
67. I. Offen-Polak *et al.*, A biomimetic twisting strategy enables efficient electrocatalytic oxidation of energy-dense hydrazine hydrate on $FeN_{2+2}C_{4+4}$ sites. *J. Am. Chem. Soc.* **147**, 31731–31740 (2025). doi: 10.1021/jacs.5c08450; pmid: 40839880
68. N. Bevilacqua *et al.*, Impact of catalyst layer morphology on the operation of high temperature PEM fuel cells. *J. Power Sources Adv.* **7**, 100042 (2021). doi: 10.1016/j.powera.2020.100042
69. J. Landesfeind, J. Hattendorff, A. Ehrl, W. A. Wall, H. A. Gasteiger, Tortuosity Determination of Battery Electrodes and Separators by Impedance Spectroscopy. *J. Electrochem. Soc.* **163**, A1373–A1387 (2016). doi: 10.1149/2.1141607jes
70. C. P. Liu *et al.*, Measurement of Contact Angles at Carbon Fiber-Water-Air Triple-Phase Boundaries Inside Gas Diffusion Layers Using X-ray Computed Tomography. *ACS Appl. Mater. Interfaces* **13**, 20002–20013 (2021). doi: 10.1021/acsami.1c00849; pmid: 33881835
71. P. Satjaritanun *et al.*, Observation of Preferential Pathways for Oxygen Removal through Porous Transport Layers of Polymer Electrolyte Water Electrolyzers. *iScience* **23**, 101783 (2020). doi: 10.1016/j.isci.2020.101783; pmid: 33294791
72. A. A. Wong, S. M. Rubinstein, M. J. Aziz, Direct visualization of electrochemical reactions and heterogeneous transport within porous electrodes in operando by fluorescence microscopy. *Cell Rep. Phys. Sci.* **2**, 100388 (2021). doi: 10.1016/j.xcrp.2021.100388
73. T. Cawte, A. Bazylak, Accurately predicting transport properties of porous fibrous materials by machine learning methods. *Electrochem. Sci. Adv.* **3**, e2100185 (2023). doi: 10.1002/elsa.202100185
74. P. A. García-Salaberri, I. V. Zenyuk, A general-purpose tool for modeling multifunctional thin porous media (*POREnet*): From pore network to effective property tensors. *Heliyon* **10**, e26253 (2024). doi: 10.1016/j.heliyon.2024.e26253; pmid: 38404803
75. Z. Chen, D. L. Danilov, R.-A. Eichel, P. H. L. Notten, Porous Electrode Modeling and its Applications to Li-Ion Batteries. *Adv. Energy Mater.* **12**, 2201506 (2022). doi: 10.1002/aenm.202201506
76. A. Vu, Y. Qian, A. Stein, Porous Electrode Materials for Lithium-Ion Batteries – How to Prepare Them and What Makes Them Special. *Adv. Energy Mater.* **2**, 1056–1085 (2012). doi: 10.1002/aenm.201200320
77. Y.-X. Yao, L. Xu, C. Yan, Q. Zhang, Principles and trends in extreme fast charging lithium-ion batteries. *EES Batteries* **1**, 9–22 (2025). doi: 10.1039/D4EB00011K
78. D. Bresser *et al.*, The importance of "going nano" for high power battery materials. *J. Power Sources* **219**, 217–222 (2012). doi: 10.1016/j.jpowsour.2012.07.035
79. X. Jin *et al.*, Mesoporous Single-Crystal Lithium Titanate Enabling Fast-Charging Li-Ion Batteries. *Adv. Mater.* **34**, e2109356 (2022). doi: 10.1002/adma.202109356; pmid: 35262214

80. D. Rambabu *et al.*, An Electrically Conducting Li-Ion Metal-Organic Framework. *J. Am. Chem. Soc.* **143**, 11641–11650 (2021). doi: 10.1021/jacs.1c04591; pmid: 34309388
81. P. X. Sun *et al.*, Formation of Super-Assembled TiO_x/Zn/N-Doped Carbon Inverse Opal Towards Dendrite-Free Zn Anodes. *Angew. Chem. Int. Ed.* **61**, e202115649 (2022). doi: 10.1002/anie.202115649; pmid: 34913229
82. H. Shi *et al.*, Functional Carbons Remedy the Shuttling of Polysulfides in Lithium–Sulfur Batteries: Confining, Trapping, Blocking, and Breaking up. *Adv. Funct. Mater.* **28**, 1800508 (2018). doi: 10.1002/adfm.201800508
83. S. Zhao *et al.*, SnO_2 as Advanced Anode of Alkali-Ion Batteries: Inhibiting Sn Coarsening by Crafting Robust Physical Barriers, Void Boundaries, and Heterophase Interfaces for Superior Electrochemical Reaction Reversibility. *Adv. Energy Mater.* **10**, 1902657 (2020). doi: 10.1002/aenm.201902657
84. N. Tran *et al.*, Effective Electrochemical Charge Storage in the High-Lithium Compound Li_8ZrO_6. *ACS Appl. Energy Mater.* **2**, 1274–1287 (2019). doi: 10.1021/acsaem.8b01819
85. R. Malik, D. Burch, M. Bazant, G. Ceder, Particle size dependence of the ionic diffusivity. *Nano Lett.* **10**, 4123–4127 (2010). doi: 10.1021/nl1023595; pmid: 20795627
86. L. Cui *et al.*, Species mass transfer governs the selectivity of gas diffusion electrodes toward H_2O_2 electrosynthesis. *Nat. Commun.* **15**, 10632 (2024). doi: 10.1038/s41467-024-55091-3; pmid: 39639001
87. Y. Ouyang, Z. Zhang, D. Li, J. Chen, G. Zhang, Emerging Theory, Materials, and Screening Methods: New Opportunities for Promoting Thermoelectric Performance. *Ann. Phys.* **531**, 1800437 (2019). doi: 10.1002/andp.201800437
88. H. Goldsmid, Porous Thermoelectric Materials. *Materials* **2**, 903–910 (2009). doi: 10.3390/ma2030903
89. J. Tang *et al.*, Holey silicon as an efficient thermoelectric material. *Nano Lett.* **10**, 4279–4283 (2010). doi: 10.1021/nl102931z; pmid: 20839780
90. L. Yang *et al.*, High thermoelectric figure of merit of porous Si nanowires from 300 to 700 K. *Nat. Commun.* **12**, 3926 (2021). doi: 10.1038/s41467-021-24208-3; pmid: 34168136
91. I. Sumirat, S. Ando, S. Shimamura, Theoretical consideration of the effect of porosity on thermal conductivity of porous materials. *J. Porous Mater.* **13**, 439–443 (2006). doi: 10.1007/s10934-006-8043-0
92. S. A. Hosseini, G. Romano, P. A. Greaney, Mitigating the Effect of Nanoscale Porosity on Thermoelectric Power Factor of Si. *ACS Appl. Energy Mater.* **4**, 1915–1923 (2021). doi: 10.1021/acsaem.0c02640
93. S. Xu *et al.*, Interfacial bonding enhances thermoelectric cooling in 3D-printed materials. *Science* **387**, 845–850 (2025). doi: 10.1126/science.ads0426; pmid: 39977506
94. B. Xu *et al.*, Highly Porous Thermoelectric Nanocomposites with Low Thermal Conductivity and High Figure of Merit from Large-Scale Solution-Synthesized $Bi_2Te_{2.5}Se_{0.5}$ Hollow Nanostructures. *Angew. Chem. Int. Ed.* **56**, 3546–3551 (2017). doi: 10.1002/anie.201612041; pmid: 28079961
95. X. Shi *et al.*, Polycrystalline SnSe with Extraordinary Thermoelectric Property via Nanoporous Design. *ACS Nano* **12**, 11417–11425 (2018). doi: 10.1021/acsnano.8b06387; pmid: 30335955
96. M. Kashiwagi *et al.*, Enhanced figure of merit of a porous thin film of bismuth antimony telluride. *Appl. Phys. Lett.* **98**, 023114 (2011). doi: 10.1063/1.3543852
97. Y. Q. Cao, T. J. Zhu, X. B. Zhao, X. B. Zhang, J. P. Tu, Nanostructuring and improved performance of ternary Bi-Sb-Te thermoelectric materials. *Appl. Phys., A Mater. Sci. Process.* **92**, 321–324 (2008). doi: 10.1007/s00339-008-4518-y
98. T. Chen *et al.*, Dimensionality Modulates Electrical Conductivity in Compositionally Constant One-, Two-, and Three-Dimensional Frameworks. *J. Am. Chem. Soc.* **144**, 5583–5593 (2022). doi: 10.1021/jacs.2c00614; pmid: 35290048
99. A. Jadhav, K. Gupta, P. Ninawe, N. Ballav, Imparting Multifunctionality by Utilizing Biporosity in a Zirconium-Based Metal-Organic Framework. *Angew. Chem. Int. Ed.* **59**, 2215–2219 (2020). doi: 10.1002/anie.201910625; pmid: 31642177
100. H. M. Ali *et al.*, Advances in thermal energy storage: Fundamentals and applications. *Pror. Energy Combust. Sci.* **100**, 101109 (2024). doi: 10.1016/j.pecs.2023.101109
101. H. Bao, Z. Ma, in *Storing Energy*, T. M. Letcher, Ed. (Elsevier, ed. 2, 2022), pp. 651–683. doi: 10.1016/B978-0-12-824510-1.00028-3
102. W. Aftab *et al.*, Phase change material-integrated latent heat storage systems for sustainable energy solutions. *Energy Environ. Sci.* **14**, 4268–4291 (2021). doi: 10.1039/D1EE00527H
103. P. Liu *et al.*, Aerogels Meet Phase Change Materials: Fundamentals, Advances, and Beyond. *ACS Nano* **16**, 15586–15626 (2022). doi: 10.1021/acsnano.2c05067; pmid: 36226846
104. L. Chen *et al.*, Electro- and photodriven phase change composites based on wax-infiltrated carbon nanotube sponges. *ACS Nano* **6**, 10884–10892 (2012). doi: 10.1021/nn304310n; pmid: 23157567
105. C. Yu, Y. Song, Modified Supporting Materials to Fabricate Form Stable Phase Change Material with High Thermal Energy Storage. *Molecules* **28**, 1309 (2023). doi: 10.3390/molecules28031309; pmid: 36770976
106. Y. Lu *et al.*, Core-Sheath Paraffin-Wax-Loaded Nanofibers by Electrospinning for Heat Storage. *ACS Appl. Mater. Interfaces* **10**, 12759–12767 (2018). doi: 10.1021/acsami.8b02057; pmid: 29589429
107. N. Tran, W. Zhao, F. Carlson, J. H. Davidson, A. Stein, Metal Nanoparticle–Carbon Matrix Composites with Tunable Melting Temperature as Phase-Change Materials for Thermal Energy Storage. *ACS Appl. Nano Mater.* **1**, 1894–1903 (2018). doi: 10.1021/acsanm.8b00290
108. H. Gao *et al.*, Nanoconfinement effects on thermal properties of nanoporous shape-stabilized composite PCMs: A review. *Nano Energy* **53**, 769–797 (2018). doi: 10.1016/j.nanoen.2018.09.007
109. T. Nomura *et al.*, Shape-stabilized phase change composite by impregnation of octadecane into mesoporous SiO_2. *Sol. Energy Mater. Sol. Cells* **143**, 424–429 (2015). doi: 10.1016/j.solmat.2015.07.028
110. T. Saboori, L. Zhao, M. Mesgarpour, S. Wongwises, O. Mahian, A novel macro-scale machine learning prediction based on high-fidelity CFD simulations: A case study on the pore-scale porous Trombe wall with phase change material capsulation. *J. Build. Eng.* **54**, 104505 (2022). doi: 10.1016/j.jobe.2022.104505
111. D. Eisenberg *et al.*, The evolution of hierarchical porosity in self-templated nitrogen-doped carbons and its effect on oxygen reduction electrocatalysis. *RSC Adv.* **6**, 80398–80407 (2016). doi: 10.1039/C6RA16606G
112. C. Agrafiotis, S. Tescari, M. Roeb, M. Schmücker, C. Sattler, Exploitation of thermochemical cycles based on solid oxide redox systems for thermochemical storage of solar heat. Part 3: Cobalt oxide monolithic porous structures as integrated thermochemical reactors/heat exchangers. *Sol. Energy* **114**, 459–475 (2015). doi: 10.1016/j.solener.2014.12.037
113. L. Shere, S. Trivedi, S. Roberts, A. Sciacovelli, Y. Ding, Synthesis and Characterization of Thermochemical Storage Material Combining Porous Zeolite and Inorganic Salts. *Heat Transf. Eng.* **40**, 1176–1181 (2018). doi: 10.1080/01457632.2018.1457266
114. X. Han *et al.*, Critical review of thermochemical energy storage systems based on cobalt, manganese, and copper oxides. *Renew. Sustain. Energy Rev.* **158**, 112076 (2022). doi: 10.1016/j.rser.2022.112076
115. M. Benitez-Guerrero *et al.*, Large-scale high-temperature solar energy storage using natural minerals. *Sol. Energy Mater. Sol. Cells* **168**, 14–21 (2017). doi: 10.1016/j.solmat.2017.04.013
116. Y. Xu *et al.*, High-Temperature Thermochemical Heat Storage Performance of CaO Honeycombs During CaO/$CaCO_3$ Cycles. *Energy Fuels* **35**, 16882–16893 (2021). doi: 10.1021/acs.energyfuels.1c02274
117. G. Seitz, R. Helmig, H. Class, A numerical modeling study on the influence of porosity changes during thermochemical heat storage. *Appl. Energy* **259**, 114152 (2020). doi: 10.1016/j.apenergy.2019.114152
118. M. Benitez-Guerrero, J. M. Valverde, P. E. Sanchez-Jimenez, A. Perejon, L. A. Perez-Maqueda, Multicycle activity of natural $CaCO_3$ minerals for thermochemical energy storage in Concentrated Solar Power plants. *Sol. Energy* **153**, 188–199 (2017). doi: 10.1016/j.solener.2017.05.068
119. C. Qin, H. Du, L. Liu, J. Yin, B. Feng, CO_2 Capture Performance and Attrition Property of CaO-Based Pellets Manufactured from Organometallic Calcium Precursors by Extrusion. *Energy Fuels* **28**, 329–339 (2014). doi: 10.1021/ef4012609
120. B. Li, Y. Li, H. Sun, Y. Wang, Z. Wang, Thermochemical Heat Storage Performance of CaO Pellets Fabricated by Extrusion–Spheronization under Harsh Calcination Conditions. *Energy Fuels* **34**, 6462–6473 (2020). doi: 10.1021/acs.energyfuels.0c00644
121. Y. Kim *et al.*, Ultrahigh-Porosity MgO Microparticles for Heat-Energy Storage. *Adv. Mater.* **35**, 2204775 (2023). doi: 10.1002/adma.202204775; pmid: 35819877
122. C. Agrafiotis, M. Roeb, M. Schmücker, C. Sattler, Exploitation of thermochemical cycles based on solid oxide redox systems for thermochemical storage of solar heat. Part 2: Redox oxide-coated porous ceramic structures as integrated thermochemical reactors/heat exchangers. *Sol. Energy* **114**, 440–458 (2015). doi: 10.1016/j.solener.2014.12.036
123. L. Liu *et al.*, Superior thermochemical energy storage performance of the Co_3O_4/CoO redox couple with a cubic micro-nanostructure. *J. Energy Storage* **43**, 103167 (2021). doi: 10.1016/j.est.2021.103167
124. P. Yuan, H. Xu, Z. Ning, G. Xiao, Understanding thermochemical energy storage performance of $Ba_{1-x}Sr_xCoO_{3-\delta}$ perovskite system: A computational and experimental study. *J. Energy Storage* **61**, 106695 (2023). doi: 10.1016/j.est.2023.106695
125. S. Tescari *et al.*, Experimental proof of concept of a pilot-scale thermochemical storage unit. *AIP Conf. Proc.* **1850**, 090006 (2017). doi: 10.1063/1.4984455
126. M. Otto *et al.*, Black Silicon Photovoltaics. *Adv. Opt. Mater.* **3**, 147–164 (2015). doi: 10.1002/adom.201400395
127. A. Stein, Achieving Functionality and Multifunctionality through Bulk and Interfacial Structuring of Colloidal-Crystal-Templated Materials. *Langmuir* **39**, 2890–2910 (2023). doi: 10.1021/acs.langmuir.2c03297; pmid: 36757136
128. R. Fayad, L. Halaoui, The Role of Order in the Amplification of Light-Energy Conversion in a Dye-Sensitized Solar Cell Coupled to a Photonic Crystal. *ChemPhysChem* **17**, 260–269 (2016). doi: 10.1002/cphc.201500942; pmid: 26643111
129. S. C. Warren *et al.*, Identifying champion nanostructures for solar water-splitting. *Nat. Mater.* **12**, 842–849 (2013). doi: 10.1038/nmat3684; pmid: 23832125
130. Y. B. Liu, R. Jin, J. Qiu, L. H. Liu, Spectral radiative properties of a nickel porous microstructure and magnetic polariton resonance for light trapping. *Int. J. Heat Mass Transf.* **98**, 833–844 (2016). doi: 10.1016/j.ijheatmasstransfer.2016.03.071

131. K. A. Amedome Min-Dianey, H.-C. Zhang, A. A. Brohi, H. Yu, X. Xia, Optical spectra of composite silver-porous silicon (Ag-pSi) nanostructure based periodical lattice. *Superlattices Microstruct.* **115**, 168–176 (2018). doi: 10.1016/j.spmi.2018.01.028
132. H.-W. Chen *et al.*, Numerical design of dual-scale foams to enhance radiation absorption. *AIMS Energy* **9**, 842–853 (2021). doi: 10.3934/energy.2021039
133. M. Kokkonen *et al.*, Advanced research trends in dye-sensitized solar cells. *J. Mater. Chem. A* **9**, 10527–10545 (2021). doi: 10.1039/D1TA00690H; pmid: 33996094
134. Y. Yu *et al.*, Development of lead iodide perovskite solar cells using three-dimensional titanium dioxide nanowire architectures. *ACS Nano* **9**, 564–572 (2015). doi: 10.1021/nn5058672; pmid: 25549153
135. W.-Q. Wu, F. Huang, D. Chen, Y.-B. Cheng, R. A. Caruso, Solvent-Mediated Dimension Tuning of Semiconducting Oxide Nanostructures as Efficient Charge Extraction Thin Films for Perovskite Solar Cells with Efficiency Exceeding 16%. *Adv. Energy Mater.* **6**, 1502027 (2016). doi: 10.1002/aenm.201502027
136. A. R. Pascoe *et al.*, Enhancing the Optoelectronic Performance of Perovskite Solar Cells via a Textured $CH_3NH_3PbI_3$ Morphology. *Adv. Funct. Mater.* **26**, 1278–1285 (2016). doi: 10.1002/adfm.201504190
137. S. Wu *et al.*, 2D metal-organic framework for stable perovskite solar cells with minimized lead leakage. *Nat. Nanotechnol.* **15**, 934–940 (2020). doi: 10.1038/s41565-020-0765-7; pmid: 32958933
138. X. Zhou *et al.*, Toward high-efficiency and thermally-stable perovskite solar cells: A novel metal-organic framework with active pyridyl sites replacing 4-*tert*-butylpyridine. *J. Power Sources* **473**, 228556 (2020). doi: 10.1016/j.jpowsour.2020.228556
139. X. Zhou *et al.*, Heterojunction Incorporating Perovskite and Microporous Metal-Organic Framework Nanocrystals for Efficient and Stable Solar Cells. *Nano-Micro Lett.* **12**, 80 (2020). doi: 10.1007/s40820-020-00417-1; pmid: 34138085
140. C. Li *et al.*, Mitigation of vacancy with ammonium salt-trapped ZIF-8 capsules for stable perovskite solar cells through simultaneous compensation and loss inhibition. *Nanoscale Adv.* **3**, 3554–3562 (2021). doi: 10.1039/D1NA00173F; pmid: 36133714
141. X. Hu *et al.*, Metal-organic framework-based catalysts for lithium-sulfur batteries. *Coord. Chem. Rev.* **475**, 214879 (2023). doi: 10.1016/j.ccr.2022.214879
142. P. Liu *et al.*, Polypyrrole-boosted photothermal energy storage in MOF-based phase change materials. *Interdiscip. Mater.* **2**, 423–433 (2023). doi: 10.1002/idm2.12092
143. X. Dai, S. Haussener, Non-Uniform Porous Structures and Cycling Control for Optimized Fixed-Bed Solar Thermochemical Water Splitting. *J. Sol. Energy Eng.* **144**, 030904 (2022). doi: 10.1115/1.4052960
144. S. Ackermann, M. Takacs, J. Scheffe, A. Steinfeld, Reticulated porous ceria undergoing thermochemical reduction with high-flux irradiation. *Int. J. Heat Mass Transf.* **107**, 439–449 (2017). doi: 10.1016/j.ijheatmasstransfer.2016.11.032
145. A. Haeussler, S. Abanades, A. Julbe, J. Jouannaux, B. Cartoixa, Solar thermochemical fuel production from H_2O and CO_2 splitting via two-step redox cycling of reticulated porous ceria structures integrated in a monolithic cavity-type reactor. *Energy* **201**, 117649 (2020). doi: 10.1016/j.energy.2020.117649
146. P. Furler *et al.*, Solar Thermochemical CO_2 Splitting Utilizing a Reticulated Porous Ceria Redox System. *Energy Fuels* **26**, 7051–7059 (2012). doi: 10.1021/ef3013757
147. M. Zarrouati, F. Enguehard, J. Taine, Statistical characterization of near-wall radiative properties of a statistically non-homogeneous and anisotropic porous medium. *Int. J. Heat Mass Transf.* **67**, 776–783 (2013). doi: 10.1016/j.ijheatmasstransfer.2013.08.021
148. S. Suter, A. Steinfeld, S. Haussener, Pore-level engineering of macroporous media for increased performance of solar-driven thermochemical fuel processing. *Int. J. Heat Mass Transf.* **78**, 688–698 (2014). doi: 10.1016/j.ijheatmasstransfer.2014.07.020
149. B. Liu, X.-L. Xia, X. Chen, M. Ma, Integrated-scale prediction of radiative properties of open-cell metal foam based on the pore-scale transfer and microscopic morphology of ligaments. *Int. J. Therm. Sci.* **163**, 106750 (2021). doi: 10.1016/j.ijthermalsci.2020.106750
150. J. J. Hwang. P. Y. Chen, Heat/mass transfer in porous electrodes of fuel cells. *Int. J. Heat Mass Transf.* **49**, 2315–2327 (2006). doi: 10.1016/j.ijheatmasstransfer.2005.11.021
151. C. A. Smith *et al.*, New approach for measuring interconnectivity of fission gas pores in nuclear fuels from 2D micrographs. *J. Mater. Sci.* **56**, 543–557 (2021). doi: 10.1007/s10853-020-05368-x
152. S. Aravindan, S. Jalaldeen, P. Chellapandi, N. Swaminathan, Yield behavior of porous nuclear fuel (UO_2). *Mech. Adv. Mater. Structures* **23**, 1149–1162 (2016). doi: 10.1080/15376494.2015.1059529
153. X. Wang, Z. Chen, K. Jiang, M. Chen, S. Passerini, 3D Host Design Strategies Guiding "Bottom–Up" Lithium Deposition: A Review. *Adv. Energy Mater.* **14**, 2304229 (2024). doi: 10.1002/aenm.202304229
154. J. Liu *et al.*, Coupled continuum and network model framework to study catalyst layers of polymer electrolyte fuel cells. *Int. J. Hydrogen Energy* **47**, 17749–17761 (2022). doi: 10.1016/j.ijhydene.2022.03.266
155. H. Dong, M. J. Blunt, Pore-network extraction from micro-computerized-tomography images. *Phys. Rev. E* **80**, 036307 (2009). doi: 10.1103/PhysRevE.80.036307; pmid: 19905212
156. M. Pang *et al.*, Facilitating Rapid OH^-/H_2O Transport in Anion Exchange Membranes via Ultra-Stable Heteroatom-Free Micropores. *Angew. Chem.* **137**, e202418435 (2025). doi: 10.1002/ange.202418435
157. W. J. Roth *et al.*, Current State and Perspectives of Exfoliated Zeolites. *Adv. Mater.* **36**, e2307341 (2024). doi: 10.1002/adma.202307341; pmid: 37800413
158. S. Suter, M. Cantoni, Y. K. Gaudy, S. Pokrant, S. Haussener, Linking morphology and multi-physical transport in structured photoelectrodes. *Sustain. Energy Fuels* **2**, 2661–2673 (2018). doi: 10.1039/C8SE00215K
159. A. Bilican, M. Leutzsch, W. Schmidt, Evaluating Nano-Scale Shifts: Quantifying Mesopore Shrinkage in Porous Polymers via NMR Cryoporometry. *Small* **21**, e2500343 (2025). doi: 10.1002/smll.202500343; pmid: 40181617
160. M. J. Werny, F. Meirer, B. M. Weckhuysen, Visualizing the Structure, Composition and Activity of Single Catalyst Particles for Olefin Polymerization and Polyolefin Decomposition. *Angew. Chem. Int. Ed.* **63**, e202306033 (2024). doi: 10.1002/anie.202306033; pmid: 37782261
161. Q. Qiu *et al.*, Recent Advances in Biomimetic Porous Materials for Real-World Applications. *Biomimetics* **10**, 521 (2025). doi: 10.3390/biomimetics10080521; pmid: 40862894
162. S. Ahmim, M. Almanza, A. Pasko, F. Mazaleyrat, M. LoBue, Thermal energy harvesting system based on magnetocaloric materials. *Eur. Phys. J. Appl. Phys.* **85**, 10902 (2019). doi: 10.1051/epjap/2019180284
163. M. Hassanpour Amiri, K. Asadi, How Porosity Affects the Performance of Piezoelectric Energy Harvesters and Sensors. *Adv. Phys. Res.* **2**, 2200042 (2023). doi: 10.1002/apxr.202200042
164. R. M. DuChanois, C. J. Porter, C. Violet, R. Verduzco, M. Elimelech, Membrane Materials for Selective Ion Separations at the Water-Energy Nexus. *Adv. Mater.* **33**, e2101312 (2021). doi: 10.1002/adma.202101312; pmid: 34396602
165. Y. Li *et al.*, Advanced biomimetic design strategies for porous structures promoting bone integration with additive-manufactured Ti6Al4V scaffolds. *J. Mater. Res. Technol.* **32**, 1901–1915 (2024). doi: 10.1016/j.jmrt.2024.08.040
166. L. Cao, M. Dong, X. Guo, M. Li, Recent Advances in Porous Low-*k* Materials for Integrated Circuits. *ACS Appl. Mater. Interfaces* **17**, 43983–44010 (2025). doi: 10.1021/acsami.5c07498; pmid: 40719332
167. J. Rouquerol *et al.*, Recommendations for the characterization of porous solids (Technical Report). *Pure Appl. Chem.* **66**, 1739–1758 (1994). doi: 10.1351/pac199466081739
168. R. G. Loucks, R. M. Reed, S. C. Ruppel, U. Hammes, Spectrum of pore types and networks in mudrocks and a descriptive classification for matrix-related mudrock pores. *AAPG Bull.* **96**, 1071–1098 (2012). doi: 10.1306/08171111061
169. K. Hu *et al.*, From Micropores to Macropores: Investigating Pore Characteristics of Longmaxi Shale in the Sichuan Basin. *Energy Fuels* **38**, 3961–3981 (2024). doi: 10.1021/acs.energyfuels.3c04516
170. C. Spurin *et al.*, Red noise in steady-state multiphase flow in porous media. *Water Resour. Res.* **58**, e2022WR031947 (2022). doi: 10.1029/2022WR031947
171. D. Kulkarni *et al.*, Influence of Microporous Layers on Interfacial Properties, Oxygen Flow Distribution, and Durability of Proton Exchange Membrane Water Electrolyzers. *ACS Appl. Mater. Interfaces* **15**, 48060–48071 (2023). doi: 10.1021/acsami.3c06899; pmid: 37788359

ACKNOWLEDGMENTS

We thank M. K. Mudunuru, S. Mondal, I. Salton, T. Burshtein, Y. Amouyal, Y. Edery, E. T. C. Vogt, and the members of the Structures of Strength Centre, NL, for advice and T. Hartman for graphic design. **Funding:** This work received funding from Israel Ministry of Energy grant 218-11-024, NWO Veni 19121, and NWO DEEP.NL.2019.006. **Competing interests:** The authors declare that they have no competing interests. **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

Submitted 21 January 2024; accepted 4 September 2025

10.1126/science.adn9391

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**



DEVELOPMENT

Cells whisper to each other in the ear

The organ of Corti contains a highly organized epithelium with ciliated (hairy) cells that sense sounds and transmit them to the brain. Sound intensity and pitch sensing rely on the precise alignment of alternating rows of various types of hair cells and supporting cells. Considering that each row contains 3500 hair cells, understanding how the hearing epithelium acquires its precise pattern is a long-standing challenge. García-Gómez *et al.* now show that extensive local communication between cells underlies this precision. Messages between cells are multiple and diverse, including instructions about alternative fate choices, cell type induction, cell density, and cell attraction. These mechanisms explain patterning robustness and help in understanding congenital and acquired sensorineural hearing deficits. —Miguel Torres Sci. Adv. (2025) 10.1126/sciadv.adz3944

Ciliated cells (yellow and red) in the organ of Corti, seen in a colored scanning electron microscopy image

IMAGE: G. BREDERG/SCIENCE SOURCE

MOLECULAR BIOLOGY

Boosting telomerase at chromosomes' ends

Human replication protein A (RPA), long recognized as an essential factor in DNA replication and repair, has now been found to play a critical role at chromosome ends. Agrawal *et al*. demonstrated that RPA directly enhances the processivity of telomerase, the enzyme responsible for extending telomeres. By boosting telomerase's ability to add repeated DNA sequences, RPA proves vital for maintaining telomere length and genomic stability. This repurposing of a fundamental single-stranded DNA-binding complex provides fresh insights into telomere regulation and has implications for aging, cancer, and telomere-related disorders. —Di Jiang

Science p. 495, 10.1126/science.ads5297

POLYMER RECYCLING

Uncovering a urethanase

There have been major advances recently in the discovery and engineering of hydrolase enzymes for degradation of ester linkages in polymers such as polyethylene terephthalate. However, more resistant chemical linkages, such as those found in urethane polymers, present additional challenges. Chen *et al*. developed a computational pipeline for screening enzymes with potential activity and identified a highly active urethanase. The authors report a kilogram-scale chemoenzymatic depolymerization of commercial polyurethane foam and show enzymatic degradation of *N*-aryl carbamates with nearly complete conversion within 8 hours. This discovery framework may be useful in identifying enzymes for the chemical and polymer recycling industry. —Michael A. Funk

Science p. 503, 10.1126/science.adw4487

COMPARATIVE COGNITION

Rational chimps

The ability of humans to think rationally and weigh the evidence when making a choice is well known. Such decision-making requires a metacognitive process in which an individual can evaluate an overall set of evidence and make the best supported choice. Whether other animals can also do this has been unknown. Schleihauf *et al*. tested whether our closest relatives, chimpanzees, were able to evaluate weak and strong evidence regarding the location of a food reward (see the Perspective by Hare). They found that the chimps correctly inferred the most rational location based on the strength of the evidence that they received about the reward's location. —Sacha Vignieri

Science p. 521, 10.1126/science.adq5229; see also p. 444, 10.1126/science.aeb7565

EPIDEMIOLOGY

COVID-19's traveling companions

The massive range and costs of interventions undertaken to limit the harms caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) had many side effects. Using surveillance data from the Centers for Disease Control and Prevention for notifiable infectious diseases, Brett and Rohani made a comprehensive statistical analysis of the indirect impact of the pandemic on a range of infections occurring in the US (see the Perspective by Durrheim). Although several endemic infections have rebounded to prepandemic levels of notification since 2020, the overall burden from infectious disease has been reduced. —Caroline Ash

Science p. 510, 10.1126/science.adw4964; see also p. 442, 10.1126/science.aeb5722

NEUROINFLAMMATION

Unsilencing neurons with BTKi's

Bruton's tyrosine kinase inhibitors (BTKi's) suppress overreactive B cell signaling and could potentially curtail disease progression in multiple sclerosis. To investigate the signaling pathways involved in B cell–mediated neuronal impairment and the effects of BTKi's, Shin *et al.* performed pharmacological experiments in mouse and human coculture models. They demonstrated that lymphotoxin-α (LTα) released from proinflammatory B cells triggered necroptotic signaling in neurons, leading to neuronal depolarization and silencing. BTKi's mitigated these effects by inhibiting LTα release. Patient data demonstrating that LTα correlated with a neuronal injury marker in cerebrospinal fluid suggested that the signaling pathways characterized in vitro could be relevant for the human condition. —Daniela Neuhofer

Sci. Transl. Med. (2025) 10.1126/scitranslmed.adx2652

IMMUNOMETABOLISM

Organelle cross-talk in lymphocytes

Lymphocytes, including regulatory T (T_{reg}) cells, tune their metabolism to support their immunological functions. However, the molecular events controlling T_{reg} cell metabolic reprogramming during inflammation are not fully understood. Using a mouse model of acute inflammation, Saravia *et al.* found that activated T_{reg} cells undergo mitochondrial metabolic reprogramming dependent on Opa1, a protein that controls mitochondrial dynamics from the inner mitochondrial membrane. Loss of Opa1 triggered T_{reg} cell mitochondrial bioenergetic stress, severe inflammation, and aberrant activation of transcription factor EB (TFEB), which led to excessive lysosomal biogenesis. These findings identify a pathway of interorganelle communication between mitochondria and lysosomes that supports T_{reg} cell–suppressive functions during inflammation. —Claire Olingy

Sci. Immunol. (2025) 10.1126/sciimmunol.ads9456

GUT MICROBIOTA

Protection by sorbitol digestion

Excess dietary intake of fructose can lead to steatotic liver disease. In zebrafish depleted of gut microbiota, Jackstadt *et al.* found that dietary glucose was converted to sorbitol in the intestine, and was then converted to a fructose intermediate in the liver, resulting in hepatic steatosis. This effect was mimicked by high levels of exogenous sorbitol and attenuated by intestinal repopulation with sorbitol-degrading bacteria. These results suggest that excessive dietary intake of sorbitol, which is used as a sugar substitute, may increase the risk of steatotic liver disease. —Wei Wong

Sci. Signal. (2025) 10.1126/scisignal.adt3549

IN OTHER JOURNALS

Edited by **Corinne Simonti** and **Jesse Smith**

SEPSIS

A clearer sepsis diagnosis

Sepsis is a deadly condition caused by severe infection and the host response to it, leading to acute dysfunction of multiple organs. It is difficult to predict and it progresses rapidly, so the standard of care is to urgently start broad antibiotic coverage until the diagnosis is ascertained, which results in overtreatment with antibiotics and the spread of drug-resistant bacteria. By using machine learning to evaluate a set of mRNAs in the blood of patients presenting for emergency care, Liesenfeld *et al.* were able to identify those with incipient sepsis and distinguish bacterial infections from viral ones that would not need antibiotic treatment, all with greater accuracy than currently available laboratory testing. —Yevgeniya Nusinovich

Nat. Med. (2025) 10.1038/s41591-025-03933-y

BEHAVIOR

Seeking food

Animals facing food scarcity and caloric deficit need to modulate their behavior to maximize the chances of finding food. Kamath *et al.* studied the brain mechanisms underlying hunger-induced behavioral changes in mice. Hunger enhanced exploratory behavior and modulated risk assessment such that hungry animals were willing to spend more time and collect more food in a threatening environment compared with satiated animals. Mechanistically, Agouti-related peptide (AgRP)–expressing neurons in the arcuate nucleus of the hypothalamus reduced dopamine release in the tail of the striatum, thus increasing exploratory behavior. These results show that dopaminergic signaling is responsible for translating an internal state in behavioral changes. —Mattia Maroso

Neuron (2025) 10.1016/j.neuron.2025.09.009

IMMUNOLOGY

Interfering with an interferon dogma

Type I and III interferons (IFN-I and IFN-III) are small proteins that signal the presence of pathogens and trigger the transcription of IFN-stimulated genes (ISGs). Ngo *et al.* used genetic tools to deplete IFN-I/IFN-III–producing plasmacytoid dendritic cells (pDCs) in mice and study their contribution to host defense. In the absence of pDCs, mice could still induce ISGs through IFN-I/IFN-III production by other cell



IMAGE: NASA, ESA, STSCI; IMAGE PROCESSING: ALYSSA PAGAN (STSCI)

WWW.AVXHM.SE



GALAXIES

Simulated flybys form galaxy bars

Some spiral galaxies, including the Milky Way, contain a galactic bar structure, a linear arrangement of stars that extends between the spiral arms through the galaxy center. Galactic dynamics studies have shown how bars can be stable and self-sustaining, but it is less clear how they initially form. Łokas has investigated barred galaxies in cosmological simulations, finding multiple cases in which a bar structure formed immediately after a galaxy experienced a close flyby with another more massive galaxy. These results demonstrate a mechanism for the rapid formation of at least some bars, particularly in dense galaxy cluster environments where flybys are common. —Keith T. Smith

Astron. Astrophys. (2025)
10.1051/0004-6361/202555503

The nearby galaxy NGC 5335 has a prominent central bar structure.

types and control the dissemination of mouse cytomegalovirus. In addition, mice that lacked pDCs showed less lung tissue damage and improved survival after severe influenza virus infections. This suggests that the ability of pDCs to produce IFN-I/IFN-III is not essential for animals to mount an effective immune response and might even be detrimental in some circumstances. —Sarah H. Ross

Nat. Immunol. (2025)
10.1038/s41590-025-02288-3

CONDENSED MATTER

Electrons dancing on liquid helium

A thin layer of liquid helium provides a pristine environment for electrons trapped on the surface. Being isolated from external noise, these single electrons can act as spin or charge qubits with long coherence times. Castoria *et al.* describe a device that allows the detection and control of single electrons trapped on helium at temperatures around 1 K. Combining a superconducting coplanar waveguide with an electron trap region, the authors demonstrated the ability to load and unload the trap with single electrons from an adjoining electron reservoir. Such single-electron sensing and control will be useful for the development of quantum technologies based on electrons on helium. —Ian S. Osborne

Phys. Rev. X (2025)
10.1103/vcl7-73ms

PLASTIC RECYCLING

Deconstructing polyurethane

Polyurethane, which is used in a wide variety of foams and spongy materials, is currently difficult to recycle. Part of the trouble is that it is constructed from two types of building blocks, respectively bearing oxygen- and nitrogen-linking groups, and most breakdown methods tend to cleanly recover just one of them. Hosgor *et al.* report that heating the foams in diethyl carbonate, which functions as both a reagent and solvent, together with zinc acetate as a catalyst, can deliver both the polyol (oxygen) and dicarbamate (nitrogen) components in 90 and 70% of respective yields. —Jake S. Yeston

Green Chem. (2025)
10.1039/D5GC02533H

NEUROSCIENCE

Profiling individual hippocampal ripples

Ripples are transient neural network events characterized by high-frequency oscillations in the local field potentials of hippocampal area CA1. Hippocampal ripples have been linked to a variety of brain functions, but they have typically been treated as homogeneous, raising the question of how they can support such diverse functions. Castelli *et al.* investigated local field potentials underlying individual ripples across the CA1 layers. They found two ripple profiles that exhibited distinct laminar currents across the CA1 somatodendritic axis and differentially recruited hippocampal neurons during periods of sleep and rest. This indicates that there may be different types of ripples with distinct population activity levels, structural organization, and neuronal content serving different brain functions. —Peter Stern

Neuron (2025)
10.1016/j.neuron.2025.09.012

Edited by **Michael Funk**

MATERIALS SCIENCE

Channeling the flow

The transport of electrons, ions, or molecules through porous materials is important for a wide range of processes and occurs across a large number of length scales. Many of these processes are tied to energy technologies, whether related to the extraction of fuel from deep underground or to the transport of charge within a solid electrolyte. Farber *et al*. reviewed research on the connections between a material's porosity and its use for energy-related technologies, including areas of commonality across length scales that can manifest as general models that may have broad relevance. —Marc S. Lavine

Science p. 484, 10.1126/science.adn9391

SOLAR CELLS

Reducing recombination in tandem cells

A high-dipole-moment passivating layer can increase electron concentration across an entire intrinsic perovskite absorber and increase efficiency in fully textured perovskite–silicon tandem solar cells. Er-Raji *et al*. showed that 1,3-diaminopropane dihydroiodide creates a dipole at the C_{60} electron-transfer layer that reduces both the energy conduction band offset and interfacial carrier recombination. The authors achieved a power conversion efficiency of 33.1% and demonstrated high stability after 2 months of outdoor stability testing along the Red Sea coast. —Phil Szuromi

Science p. 488, 10.1126/science.adx1745

MOLECULAR BIOLOGY

Regulating RAD51 equilibrium

Repair of DNA double-stranded breaks (DSBs) through BRCA2-mediated homologous recombination (HR) is critical for cellular viability and tumor suppression. Traditionally, loss of BRCA2 was thought to inhibit HR by preventing the loading of the RAD51 recombinase onto resected DSBs. However, Kuthethur *et al*. demonstrated that depletion of the AAA^+ ATPase FIGNL1 in BRCA2-deficient cells restored RAD51 loading and rescued HR defects. This suggests that recombination failure in BRCA2-deficient cells is primarily due to excessive removal of RAD51 by FIGNL1 rather than to defective RAD51 loading per se. Furthermore, the researchers identified the MMS22L-TONSL complex as being essential for HR in BRCA2/FIGNL1 double-deficient cells. These findings define a regulatory mechanism controlling RAD51 equilibrium during recombination and have implications for understanding chemoresistance in BRCA2-mutant cancers. —Di Jiang

Science p. 489, 10.1126/science.adt1210

PROTEIN CHAPERONES

Building tubulin

Tubulin is a heterodimeric protein that is the basic building block for microtubules, a key component of the eukaryotic cytoskeleton. Multiple chaperones are required to put α- and β-tubulin monomers together into the functional heterodimer, and these chaperones also facilitate recycling through nucleotide exchange. Seong *et al*. determined structures of the human tubulin-binding cofactors at six stages that reveal how these proteins bind to a β-tubulin monomer, promote heterodimer formation, and control nucleotide hydrolysis and release of the matured tubulin. These structures provide a detailed view into the complex process of assembling the basic units of a cell's cytoskeleton. —Michael A. Funk

Science p. 490, 10.1126/science.ady2708

GREENHOUSE GASES

Cutting emissions will cut costs

Methane is a key greenhouse gas, and its emissions create large economic costs because of their effects on climate. Stoerk *et al*. calculated the economic value of methane emission abatement efforts, finding that reasonable actions could reduce climate-related damages in 2050 by more than a trillion dollars a year while costing only one sixth of the damages they would avoid. Concerted methane abatement action also would reduce the risks associated with climate-tipping behaviors and extreme climate damage scenarios. —Jesse Smith

Science p. 491, 10.1126/science.adu7392

IMMUNOLOGY

Gut neurons in sickness and in health

Enteric neurons of the intestine play a major role in regulating intestinal physiology. Tan *et al*. characterized the effect of microbiota and inflammation on the enteric nervous system in mice. The authors integrated single-cell transcriptomics across three intestinal regions with in vivo functional perturbation to uncover how microbial cues and immune signals reshape inhibitory motor neuron states and sensory neuron plasticity. Using single-cell RNA sequencing and genetic perturbations, they identified regulators of motor neuron states and intestinal gastric time. In addition to providing biological insights into the response of the enteric nervous system to various perturbations, the results are a valuable resource that will serve as an important benchmark for many future discoveries. —Mattia Maroso

Science p. 492, 10.1126/science.adr3545

NEUROSCIENCE

Linking mutations and neurodegeneration

Repetitive head injury can lead to the development of chronic traumatic encephalopathy (CTE), a neurodegenerative disease with unclear pathogenesis. Dong *et al*. applied single-cell RNA sequencing to brain samples from patients with CTE and identified increased somatic single-nucleotide variants and small insertions and deletions in CTE neurons compared with repetitive head injury without CTE or with neurotypical controls. These alterations resembled similar patterns seen in patients with Alzheimer's disease, suggesting that the same genetic changes might contribute to neurodegeneration in CTE. —Mattia Maroso

Science p. 493, 10.1126/science.adu1351

CATALYSIS

More olefins and less CO_2 from syngas

Two different strategies can produce olefins from synthesis gas (syngas, a mixture of carbon monoxide and hydrogen) with fewer CO_2 by-products over iron-based catalysts (see the Perspective by Saeys). Cai *et al*. fed trace amounts of bromomethane with syngas over iron-based catalysts. Surface-bound bromine interacted with iron active sites to inhibit water dissociation, carbon monoxide and oxygen atom recombination, and olefin hydrogenation, and enabled near-zero CO_2

production and high α-olefin selectivity. In another study, Gao *et al*. found that a sodium-promoted $FeC_x@Fe_3O_4$ core-shell nanoparticle catalyst could couple water-gas shift and syngas-to-olefins reactions in situ. Starting from syngas with low hydrogen/carbon monoxide ratios, the authors achieved high olefin selectivity and hydrocarbon yield, along with a reduction in CO_2 and water by-products that led to high hydrogen atom economy. —Phil Szuromi

Science p. 516, 10.1126/science.aea0774; p. 494, 10.1126/science.aea1655
see also p. 443, 10.1126/science.aec1997

EVOLUTION

Why so blue?

Parasitic cuckoos face strong selective pressure for their eggs and offspring to match their hosts, yet each species parasitizes multiple hosts. Merondun *et al*. examined the genetic basis of egg coloration in the common cuckoo (*Cuculus canorus*) and the oriental cuckoo (*C. optatus*). In *C. canorus*, egg coloration was most strongly influenced by maternal variation across morphs, and all 116 single-nucleotide polymorphisms that associated with a blue egg morph located to the female-denoting W chromosome. Males mate opportunistically in *C. canorus*, potentially explaining the benefit of these traits being closely linked in the matriline. By contrast, this trait appears to be highly polygenic and autosomal in *C. optatus*, demonstrating the myriad ways that evolution can respond to similar pressures (see the Perspective by Sorenson and Spottiswoode). —Corinne Simonti

Science p. 527, 10.1126/science.adt9355;
see also p. 440, 10.1126/science.aec1973

RESEARCH ARTICLE SUMMARY

SOLAR CELLS

Electron accumulation across the perovskite layer enhances tandem solar cells with textured silicon

Oussama Er-raji* *et al.*

Full article and list of author affiliations: https://doi.org/10.1126/science.adx1745

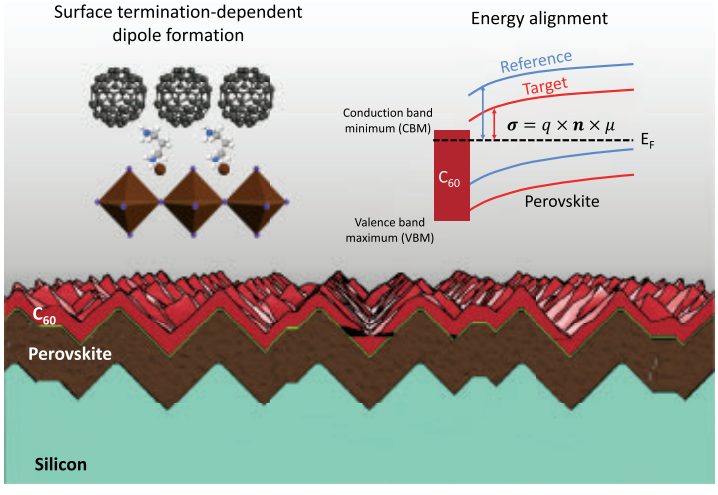
INTRODUCTION: To realize terawatt-scale photovoltaic deployment, achieving both resource and cost efficiency are imperative. Perovskite/silicon tandem solar cells present a compelling solution by surpassing the power conversion efficiency (PCE) of single-junction silicon solar cells, with only marginal additional production costs. Today, this technology achieves up to 34.85% PCE, establishing it as the most efficient dual junction. Moving toward commercialization, the adoption of industry-compatible fabrication processes and solar cell architectures becomes increasingly essential. In this context, monolithic fully textured perovskite/silicon tandem solar cells leverage the use of textured silicon substrates with industry-standard pyramid heights (>1 μm), which does not necessitate modifications to existing silicon manufacturing lines. The perovskite top cell is conformally grown on the large silicon pyramids via the evaporation–spin-coating route for the perovskite absorber, resulting in a structured perovskite/air interface that minimizes reflection losses. Despite its theoretically high PCE, substantial photovoltage and charge-transport losses limit the PCE to <32%, primarily because of inadequate interface passivation at the perovskite/electron transport layer (C_{60}) junction.

RATIONALE: Building on an optoelectrical simulation study, we identified two approaches to mitigate photovoltage losses at the perovskite/C_{60} interface: (i) chemical passivation strategies aimed at lowering surface recombination velocity, and (ii) work function engineering strategies designed to modulate charge carrier concentrations at the contact interface (i.e., field-effect passivation). Focusing on the latter, we proposed that—unlike traditional semiconductors such as silicon—owing to the intrinsic property of the perovskite film, not only the interface properties are improved by majority charge carrier (electron) accumulation at the interface. Instead, the induced electron accumulation region extends across the entire perovskite absorber, raising the electron concentration, and thus absorber conductivity. Consequently, this strategy not only reduces photovoltage losses but also mitigates transport losses, leading to improvements in both open-circuit voltage (V_{OC}) and fill factor (FF).

RESULTS: To implement the work function engineering strategy, we deposited 1,3-diaminopropane dihydroiodide on the perovskite surface, leading to a positive dipole formation (see the figure). Optoelectronic characterization revealed a reduction in the energy offset between the perovskite conduction band minimum (CBM) and the C_{60} lowest unoccupied molecular orbital, decreasing from 180 meV in the reference to 70 meV in the passivated system. Additionally, an increase of 80 mV was observed in the implied V_{OC} of perovskite/C_{60} stacks. A FF loss analysis showed that transport losses were halved, partly owing to mitigation of the non-ohmic resistance component. Consequently, passivated tandem solar cells achieved PCEs of up to 33.1%, with a V_{OC} of 2.01 V, and enhanced outdoor stability.

CONCLUSION: Minimizing energy band offsets at the perovskite/charge transport interfaces has proven effective not only in mitigating photovoltage losses through modulation of charge carrier concentrations at the interface but also in enhancing charge transport via a deep field-effect passivation mechanism. Applied to fully textured perovskite/silicon tandems, this strategy yielded a V_{OC} improvement of 100 mV compared to previous devices with identical architecture. Notably, the results confirmed that increased perovskite surface area—stemming from the large pyramid heights in industrial silicon substrates—does not pose a fundamental barrier to effective interfacial passivation in the perovskite top cell. □

*Corresponding author. Email: oussama.er-raji@ise.fraunhofer.de (O.E.); stefaan.dewolf@kaust.edu.sa (S.D.W.); stefan.glunz@ise.fraunhofer.de (S.W.G.) Cite this article as O. Er-raji *et al.*, *Science* **390**, eadx1745 (2025). DOI: 10.1126/science.adx1745



Interface passivation at the perovskite/C_{60} electron transport layer for efficient fully textured perovskite/silicon tandem solar cells. Perovskite surface treatment with 1,3-diaminopropane dihydroiodide reveals that the termination chemistry of the perovskite critically influences the orientation of the resulting dipole. This dipole minimizes the CBM energy offset between C_{60} and the perovskite (target versus reference). This effect facilitates electron accumulation both at the perovskite/C_{60} interface, which reduces interfacial recombination, and across the entire perovskite layer, which enhances conductivity, where σ is conductivity, q is elementary charge, n is electron density, and μ is mobility.

SOLAR CELLS

Electron accumulation across the perovskite layer enhances tandem solar cells with textured silicon

Oussama Er-raji[1,2,3]*, Christoph Messmer[1,2], Rakesh R. Pradhan[4], Oliver Fischer[1,2], Vladyslav Hnapovskyi[4], Sofiia Kosar[4], Marco Marengo[4], Mathias List[5], Jared Faisst[1,6], José P. Jurado[4], Oleksandr Matiash[4], Hannu P. Pasanen[4], Adi Prasetio[4], Badri Vishal[4], Shynggys Zhumagali[4], Anil R. Pininti[4], Yashika Gupta[1,2,3], Clemens Baretzky[1], Esma Ugur[4]†, Christopher E. Petoukhoff[4], Martin Bivour[1], Erkan Aydin[4]†, Randi Azmi[4]‡, Jonas Schön[1,2], Florian Schindler[1], Martin C. Schubert[1], Udo Schwingenschlögl[4], Frédéric Laquai[4]§, Ahmed A. Said[4], Juliane Borchert[1,2,3], Patricia S. C. Schulze[1], Stefaan De Wolf[4]*, Stefan W. Glunz[1,2,3]*

Reducing charge carrier transport losses, improving selectivity, and minimizing nonradiative recombination are essential for enhancing the efficiency and stability of perovskite/silicon tandem solar cells. We used a hybrid two-step perovskite deposition method that is compatible with industry-standard textured silicon, incorporating a perovskite surface treatment based on 1,3-diaminopropane dihydroiodide. The interaction of this molecule with the perovskite surface increased the majority charge carrier concentration at the electron-selective contact, which reduced interfacial recombination. Simultaneously, this field-effect passivation increased the electron concentration across the entire intrinsic perovskite absorber, which increased conductivity and reduced transport losses. Combined, this yields high-performance, fully textured perovskite/silicon tandem solar cells, achieving a 1-sun AM1.5G conversion efficiency of 33.1% with an open-circuit voltage of 2.01 volts and an extended outdoor stability in the Red Sea Coast.

Leveraging industrially mature crystalline silicon production lines, perovskite/silicon tandem solar cells promise high power conversion efficiencies (PCEs) at an affordable cost (*1*, *2*). Recent advancements in PCEs have resulted from improvements of the tandem's light-management scheme (*3*, *4*), the introduction of perovskite bulk crystallization and defect passivation agents (*5*, *6*), the emergence of hole-selective self-assembled monolayers (*7*), and particularly the development of interfacial passivation strategies at the critical perovskite/electron transport layer (C_{60}) interface (*8*–*10*). For the latter, various strategies have been proposed such as contact displacement, chemical passivation, surface-potential homogenization, and field-effect passivation (*11*–*17*). These strategies have been shown to be effective for tandems with solution-processed perovskites. However, such tandems often require tailored silicon bottom solar cells (e.g., with additional polishing or pyramid size adaptation steps), contrasting with mainstream silicon technologies, which usually apply random pyramid silicon texturing with heights above 1 μm, i.e., structure sizes appreciably larger than the thickness of the perovskite top cell.

Fully textured perovskite/silicon tandem solar cells, based on silicon with larger texture (pyramid height > 1 μm), can offer both a more cost-effective manufacturing and a superior light-management scheme (fig. S1) (*18*–*20*). These tandems are based on conformal perovskite top-cell deposition on silicon with industry-relevant texture, through a two-step, hybrid evaporation/spin-coating method (*21*–*24*). However, for this tandem design, the implementation of effective passivation interlayers at the perovskite/C_{60} interface is challenging (*25*, *26*), as evidenced by the >100-mV open-circuit voltage (V_{OC}) difference between fully textured tandems and their record-setting counterparts (fig. S2).

We show that a large energetic conduction band offset between the perovskite and the electron transport layer not only leads to a low ratio of majority versus minority (i.e., electron versus hole) charge carrier concentration at the top-contact interface (*9*, *12*), but that this effect extends to the perovskite bulk. As a result, high interfacial recombination reduces V_{OC}, and the moderate carrier concentration in the absorber limits its conductivity, inducing charge carrier transport and selectivity losses, and thus a poor fill factor (FF) of the device. We tackled this issue by treating the perovskite surface with 1,3-diaminopropane dihydroiodide (PDAI) with high dipole moment. The selective interaction of PDAI with organic compounds on the perovskite surface created a positive dipole pointing from C_{60} toward the perovskite. This dipole reduced the energy conduction band offset, inducing electron accumulation at the C_{60} junction, which reduced interfacial carrier recombination. Additionally, it led to a so far overlooked increase of electron concentration across the entire intrinsic perovskite absorber that improved its conductivity. We fabricated fully textured perovskite/silicon tandems with high PCEs, excellent reproducibility across various research institutes, and prolonged outdoor stability under hot and humid climate conditions.

Charge carrier management

We performed an optoelectrical simulation study to elaborate the different strategies to improving the perovskite/C_{60} interface, on the basis of an experimentally validated perovskite/silicon tandem solar cell model (*27*–*29*). Two parameters for the device performance are emphasized in the schematic band diagram of the perovskite/C_{60} interface (Fig. 1A): chemical passivation quality, reflected by the surface recombination velocity $S_{0,ETL}$ (highlighted in purple, vertical dashed line), and band alignment properties, represented by the perovskite/C_{60} conduction band (CB) offset $\Delta E_{C,ETL}$ (highlighted in green, horizontal dotted line). Nonideal chemical passivation leads to carrier recombination at the interface referred to as chemical passivation losses. A large CB offset leads to a lower majority-to-minority charge carrier concentration ratio that could impose selectivity losses, reflected in a gradient in the quasi-Fermi level (highlighted in green). This loss is quantified by ΔV_{OC}, defined as a difference between internal open-circuit voltage (iV_{OC}, alternatively referred to as quasi–Fermi level splitting over charge QFLS/q) and external V_{OC}.

The model distinguishes between chemical passivation losses, which arise from the difference between actual and ideal internal voltages (Fig. 1B, purple top area), and selectivity losses (Fig. 1B, green bottom area). The simulated iV_{OC} and V_{OC} are depicted in Fig. 1B as a function of $\Delta E_{C,ETL}$. Under the assumption of ideal chemical passivation ($S_{0,ETL}$ = 0, shown as a dark purple dotted line), the iV_{OC} reached nearly 2 V [passivation limit for this specific choice of bulk and hole transport layer (HTL) properties] and remained largely independent of $\Delta E_{C,ETL}$.

Many state-of-the-art devices are presently being optimized through chemical passivation strategies, which aim to reduce $S_{0,ETL}$ (see purple

[1]Fraunhofer Institute for Solar Energy Systems ISE, Freiburg, Germany. [2]Chair for Photovoltaic Energy Conversion, Department of Sustainable Systems Engineering (INATECH), University of Freiburg, Freiburg, Germany. [3]Freiburg Center for Interactive Materials and Bioinspired Technologies (FIT), University of Freiburg, Freiburg, Germany. [4]Center for Renewable Energy and Storage Technologies (CREST), Physical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. [5]Materials Research Center Freiburg (FMF), University of Freiburg, Freiburg, Germany. [6]Institute of Physics, University of Freiburg, Freiburg, Germany. ***Corresponding author. Email: oussama.er-raji@ise.fraunhofer.de (O.E.); stefaan.dewolf@kaust.edu.sa (S.D.W.); stefan.glunz@ise.fraunhofer.de (S.W.G.)** †Present address: Department of Chemistry, Ludwig-Maximilians-Universität München (LMU), Munich, Germany. ‡Present address: School of Science and Engineering, The Chinese University of Hong Kong, Shenzen (CHUK-SZ), Guangdong, China. §Present address: Department of Chemistry, Ludwig-Maximilians-Universität München (LMU), Munich, Germany.

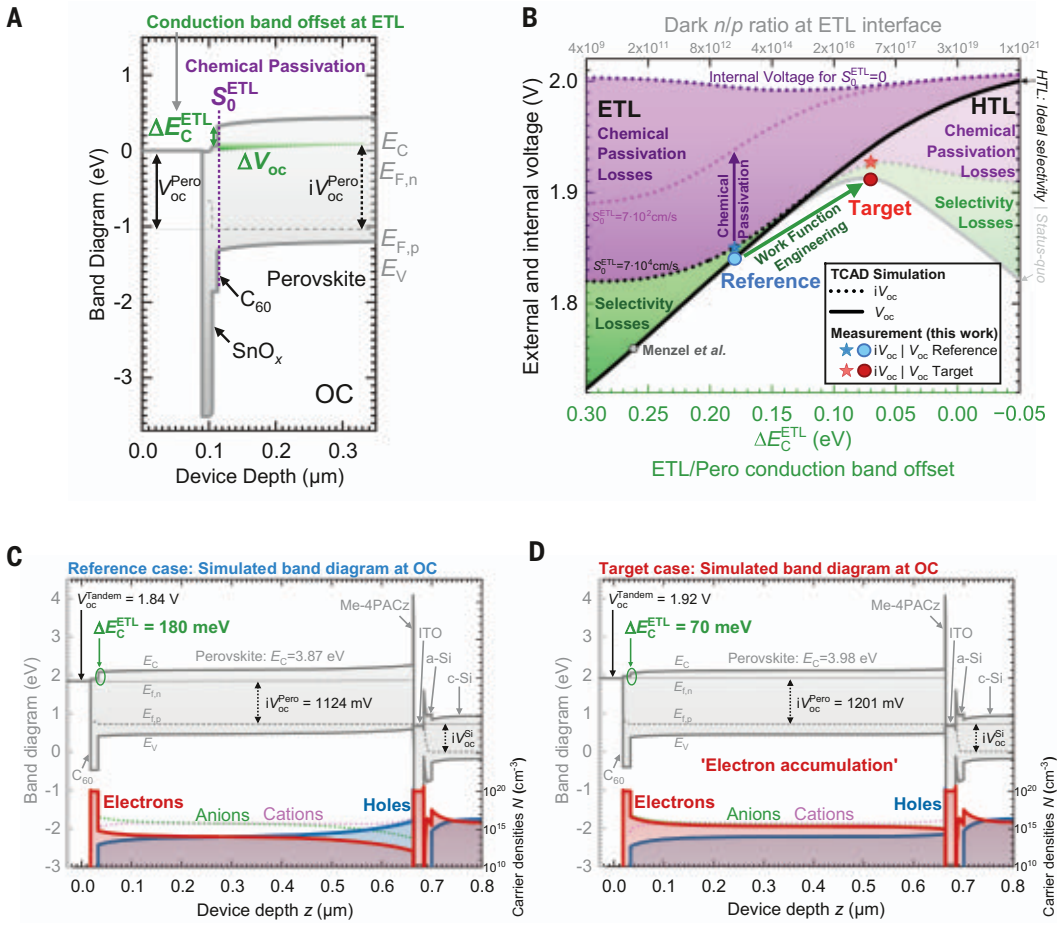


Fig. 1. Simulation-based strategies to overcome nonradiative recombination and transport losses at the perovskite's charge transport layer interfaces. (**A**) Schematic band diagram of a perovskite/C_{60} interface with chemical interface passivation S_0 (highlighted in purple) and conduction band offset $\Delta E_{C,ETL}$ (highlighted in green) under open-circuit (OC) conditions. (**B**) Internal and external voltage of a perovskite/silicon tandem solar cell as a function of the $\Delta E_{C,ETL}$, assuming ideal HTL selectivity (dotted and solid black lines) or status-quo HTL selectivity (dotted and solid gray lines). Following the work function engineering path (shown in green) reduces the interface losses as well as the selectivity losses. (**C**) Simulated band diagram under OC conditions of the experimental reference device with high $\Delta E_{C,ETL}$ and low electron density. (**D**) Simulated band diagram under OC conditions of the experiment target case (PDAI-passivated) with reduced ETL offset, resulting in electron accumulation at the ETL, and thus reduction of interfacial recombination, and increased electron concentration in the perovskite bulk (~40 times higher than that of the reference device), leading to improved electron transport.

arrow from reference case) to achieve an improvement in the vertical direction. This approach seeks to narrow the gap between iV_{OC} and the passivation limit, thereby reducing photovoltage losses (Fig. 1B and fig. S3). However, we focused on work function engineering strategies (*9*, *12*, *30*), which modulate the carrier concentrations at the contact interface by increasing the electron to hole concentration ratio (n/p) ratio to enable an improvement in the diagonal direction (see green diagonal arrow in Fig. 1B). Under dark conditions, an exponential increase in n/p is obtained for a linear increase in contact work function (fig. S4, A and B).

This approach is usually called field-effect passivation and similarly reduces photovoltage losses by reducing recombination at the interface (*12*, *31*). However, another key advantage of this strategy is that not only the interface properties are improved by electron accumulation, but because of the intrinsic properties of the perovskite bulk (*32*, *33*), the induced electron accumulation region extends across the entire perovskite absorber. Given the relation $\sigma_n = q \times n \times \mu_n$, where σ is conductivity, q is elementary charge, n is electron density, and μ is mobility, this approach increases the conductivity of the absorber, contributing to a reduction of transport losses, and is thus a primary driver for *FF* improvement of the device.

To showcase how work function engineering (in this case by engineering the perovskite's CB energy) regulates charge carrier concentrations, we simulated band diagrams of the perovskite subcell (Fig. 1, C and D). For the band diagram of the reference device in this work (without PDAI), the measured perovskite CB energy of 3.87 eV was used, leading to a high $\Delta E_{C,ETL}$ of 180 meV. Thus, only a moderate electron concentration of 6.3×10^{14} cm^{-3} was obtained at the ETL, under open-circuit condition and at 1-sun illumination (Fig. 1C).

By contrast, for the band diagram of the target device (with PDAI), we changed the perovskite CB energy to the value of 3.98 eV by increasing the electron affinity, while conserving a similar bandgap and doping concentration. These conditions resulted in a reduced $\Delta E_{C,ETL}$ of 70 meV and led to a higher electron concentration at the ETL of 4.4×10^{16} cm^{-3} (Fig. 1D).

Moreover, the simulation model revealed that this high n/p ratio was sustained both under 1-sun illumination at OC and maximum power point (MPP) conditions (fig. S4C). This electron accumulation at the ETL interface led to field-effect passivation that increased the iV_{OC} within the perovskite from 1.12 to 1.20 V (for both cases, the same surface recombination velocity $S_{0,ETL} = 7 \times 10^4$ cm/s was used), which raised the external voltage of the tandem device from 1.84 to 1.92 V (Fig. 1, C and D). The work function engineering at the electron contact increased the average electron density in the entire perovskite bulk from 1×10^{14} cm^{-3} for the simulated reference device to 4×10^{15} cm^{-3}

in the target device, enhancing the electron conductivity, and thus electron transport, in the whole perovskite bulk toward the ETL (Fig. 1D).

For further ETL optimization, we also considered the HTL. The solid black line in Fig. 1B shows the V_{OC} for a nonlimiting selectivity of the HTL (and current properties of the ETL), where increased performance is expected going toward $\Delta E_{C,ETL} \leq 0$ (right side of Fig. 1B). By contrast, the dotted and solid gray lines show the iV_{OC} and V_{OC} for an HTL selectivity expected for the experimentally used Me-4PACz. Work function engineering at the ETL side is not independent of the HTL side because of the high Debye length of the intrinsic perovskite bulk. Instead, iV_{OC} is expected to decline for increasing electron concentrations in the bulk, and HTL selectivity losses in the bulk would be expected to increase. Thus, parallel engineering at the HTL side at advanced stages of ETL optimization is necessary.

Molecule-to-surface interaction

We executed the work function engineering strategy by introducing PDAI molecules at the perovskite/C_{60} interface through a wet-chemical route (spin-coating) that was scalable (dip-coating), as detailed in the materials and methods section (perovskite composition $Cs_{0.13}FA_{0.87}Pb(I_{0.74}Br_{0.26})_3$). X-ray photoelectron spectroscopy (XPS) revealed signatures of PDAI molecules on the perovskite film surface after treatment (fig. S5). To exclude the formation of low-dimensional perovskite phases that have been reported to occur upon interaction of large organic molecules (*34*), such as the one used in this study, we further performed structural characterization. X-ray diffraction (XRD), grazing-incidence wide-angle x-ray scattering (GIWAXS), and scanning transmission electron microscopy (STEM) revealed the absence of such phases in our system (figs. S6 and S7), potentially because of the low concentration of molecules used. Additionally, no effect was observed on the perovskite morphology (fig. S8).

We investigated the band alignment at the perovskite/C_{60} interface with and without PDAI modification by performing ultraviolet photoelectron spectroscopy (UPS). The PDAI-passivated perovskite film showed a 270-meV increase in the work function (W_f) compared to the reference (from 4.49 to 4.76 eV, Fig. 2A), consistent with results from amplitude-modulated Kelvin probe force microscopy (AM-KFPM, fig. S9). Changes in W_f of a semiconductor can be caused by a shift in the Fermi level (E_F) position within the bandgap (through doping or charge distribution), by a surface dipole, or both (*9*, *12*). We distinguished these effects by measuring the ionization potential (IP); whereby a change of W_f and IP by a comparable amount would suggest the formation of a dipole whereas an individual change in the W_f with a constant IP would signify a change in E_F within the bandgap (*9*, *12*). The IP in the reference perovskite was 5.54 eV versus 5.65 eV for the passivated perovskite film (Fig. 2C), indicating the formation of a dipole. By constructing energy-level diagrams of reference and target samples (Fig. 2C) and comparing their energy-level alignment with C_{60} (in particular, the alignment of conduction bands of perovskite and lowest unoccupied molecular orbital of C_{60}), we concluded that the dipole formation in the target sample led to a minimized energetic offset at the interface with C_{60}, from 180 meV (for reference) to 70 meV (for target).

We conducted density functional theory (DFT) calculations to understand the dipole formation and orientation. PDAI adopts a symmetric bent structure, and the obtained electrostatic potential map reveals pronounced regions of positive and negative values, resulting in a dipole moment of 8.27 D (Fig. 2D). The PDAI-perovskite interaction was examined for both inorganic (PbI_2) and organic (methylammonium iodide, MAI) surface terminations of a prototypical $MAPbI_3$ system (Fig. 2, E and F). The W_f calculated for the PbI_2-terminated surface was 5.95 eV versus 4.23 eV for the MAI-terminated surface (Fig. 2G), consistent with reported values (*36*). The I atom of PDAI bonded with a surface Pb atom of the PbI_2-terminated surface with an interaction energy of –1.98 eV (Fig. 2E), whereas the interaction energy between PDAI and the MAI-terminated surface was 0.81 eV (Fig. 2F). Accordingly, we found larger charge transfers for PDAI on the PbI_2-terminated surface (Fig. 2E) than on the MAI-terminated surface (Fig. 2F). The interaction of PDAI with undercoordinated Pb atoms of the inorganic-terminated surface induced a dipole moment of –1.40 D (a negative dipole) along the out-of-plane direction. By contrast, the interaction of PDAI with the organic-terminated surface induced a dipole moment of +0.27 D (positive dipole) along the out-of-plane direction. Consequently, the dipole formation reduced the W_f to 4.21 eV for the PbI_2-terminated surface and increased it to 4.78 eV for the MAI-terminated surface (Fig. 2G and fig. S10) (*36*).

The two-step hybrid method used in this study involved a heterogeneous solid-liquid reaction pathway for the conformal formation of thin films on textured silicon with large pyramid size and yields an organic-rich perovskite surface (*37*–*39*). An organic-rich surface termination would lead to a positive dipole and a subsequent increase in the W_f as revealed through DFT calculations (Fig. 2, F and G) and as was observed experimentally with UPS and AM-KPFM (Fig. 2A and fig. S9, C and D).

The DFT results exhibited slight deviations from previous reports that examined diammonium-based passivation molecules, which consistently reported a reduction of the W_f (*8*, *15*, *16*, *40*–*43*). We attributed this discrepancy to the processing method of the perovskite absorber. The previous studies predominantly used a one-step spin-coating method, known to produce inorganic PbI_2-rich films (*44*–*46*), which result in the formation of a negative dipole and consequently a lower W_f, as confirmed by our DFT results (Fig. 2G). These findings underscore the importance of the surface termination and the specific molecule-to-surface interactions in determining the dipole orientation for similarly applied passivation molecules.

Surface passivation and conductivity

To examine the photophysical properties of perovskite films with passivation, we used hyperspectral photoluminescence (PL) imaging for generating iV_{OC} images of films with and without C_{60} (Fig. 3A). The reference perovskite iV_{OC} image displayed spatial fluctuations with a standard deviation (σ) of 38 mV and an average of 1.23 V. By contrast, with PDAI treatment, σ was lowered to 26 mV and the average iV_{OC} increased by 20 mV to reach 1.25 V, demonstrating a slight chemical passivation nature of the molecule. The improved spatial homogeneity with PDAI treatment was also observed through AM-KPFM that revealed a more uniform contact potential distribution (fig. S9), consistent with a previous study (*15*). After C_{60} deposition, the iV_{OC} of the reference stack dropped by 80 mV to 1.15 V and showed increased broadening (σ = 65 mV). By contrast, the iV_{OC} of the PDAI-passivated stack was preserved at 1.23 V with a considerably narrower distribution (σ = 17 mV) (*47*).

We performed time-resolved PL measurements (Fig. 3B), from which a differential lifetime analysis was performed to decouple the competing processes of interfacial charge transfer and interfacial charge recombination (Fig. 3C) (*7*, *48*). The reference sample with C_{60} had a differential lifetime that reached a plateau (or slowest decay component) at ~1 μs, which was reduced by about one order of magnitude compared to the reference sample without C_{60}, demonstrating that interfacial recombination dominated the dynamics. In comparison, the PDAI-treated stack with C_{60} exhibited a differential lifetime plateau of ~10 μs and that without C_{60}, a differential lifetime plateau of ~13 μs, indicating reduced interfacial recombination in the target stack. We note that the differential lifetime plateau is not the effective carrier lifetime, defined as the weighted-average carrier lifetime, which in all cases was less than 100 ns.

To investigate the carrier dynamics and non-emissive trap state signals on the nanosecond timescale in more detail, we conducted transient absorption (TA) (Fig. 3D) and transient reflectance (TR)

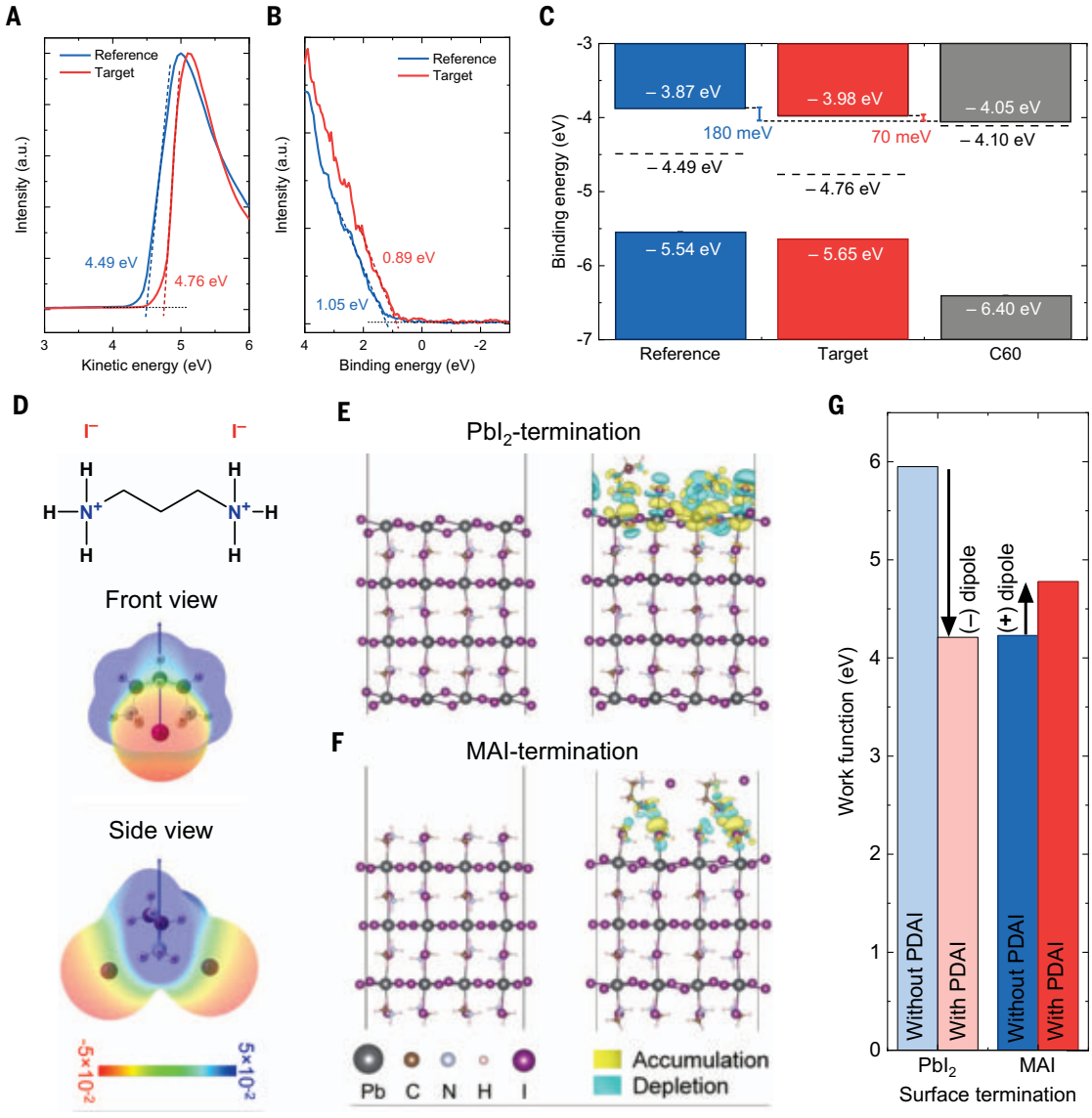


Fig. 2. Energetics and surface chemistry. (**A** and **B**) Secondary electron cutoff (A) and valence band maximum edge (B) regions of reference and target films from UPS measurements. (**C**) Energy band diagram of reference and target films compared with C_{60}. The energy levels of C_{60} are taken from (*15*,*35*). (**D**) Chemical structure and electrostatic potential map of the PDAI molecule. (**E** and **F**) Optimized structure of PbI_2-terminated $MAPbI_3$ (where MA is methylammonium) (E) and MAI-terminated $MAPbI_3$ (F) with and without PDAI. The induced charge density differences are indicated. (**G**) Work function calculated for PbI_2-terminated MAPbI3 and MAI-terminated $MAPbI_3$ with and without PDAI. A positive dipole points from C_{60} toward the perovskite (electronegative side toward C_{60}, electropositive side toward perovskite). A negative dipole points from the perovskite toward C_{60} (electronegative side toward perovskite, electropositive side toward C_{60}).

(fig. S11) spectroscopies. The bandgap bleach signal from the reference showed a long-lived state, which we attributed to charge carriers trapped in shallow traps, which were no longer present after the PDAI treatment. Similar to the PL lifetimes, the weight-averaged carrier lifetime of the TA bandgap bleach of the reference/C_{60} stack increased from 27 to 53 ns upon PDAI treatment, indicating reduced interfacial recombination as the main reason for the improved V_{OC}.

As charge transport depends on the resistivity across the interface and within the absorber, we performed suns-V_{OC} measurements on representative reference and PDAI-passivated tandem solar cells (*49*). As this measurement is carried out under open-circuit conditions, in which charge extraction is avoided, the obtained pseudo-*J*-*V* curve is considered free of transport losses and thus enables a quantification of the lumped series resistance (R_S). We found that, whereas the determined pseudo-FF (pFF) remained constant at 85%, the pFF-FF difference decreased from 6% in the reference device to 3% in the PDAI-passivated device (Fig. 3, E and F). As a result, a noticeable drop was calculated in R_S at maximum power point (MPP) from 6 to 3 ohm cm^2. These results were reproducible across different samples as shown in fig. S12 and were equally observed in optoelectrical simulations (fig. S13). Complementing the transport loss analysis, we evaluated R_S at different operating voltages using the pseudo-*J*-*V* and *J*-*V* curves (fig. S14). Both reference and target devices showed a slope in the resultant R_S versus *V* curve, associated with the presence of a non-ohmic contribution besides the classical ohmic component of R_S. Notably, the target device exhibited a reduced slope, underlining that changes in contact resistivity induced by PDAI are minor, and that PDAI mainly reduces the non-ohmic contribution of R_S through modulation

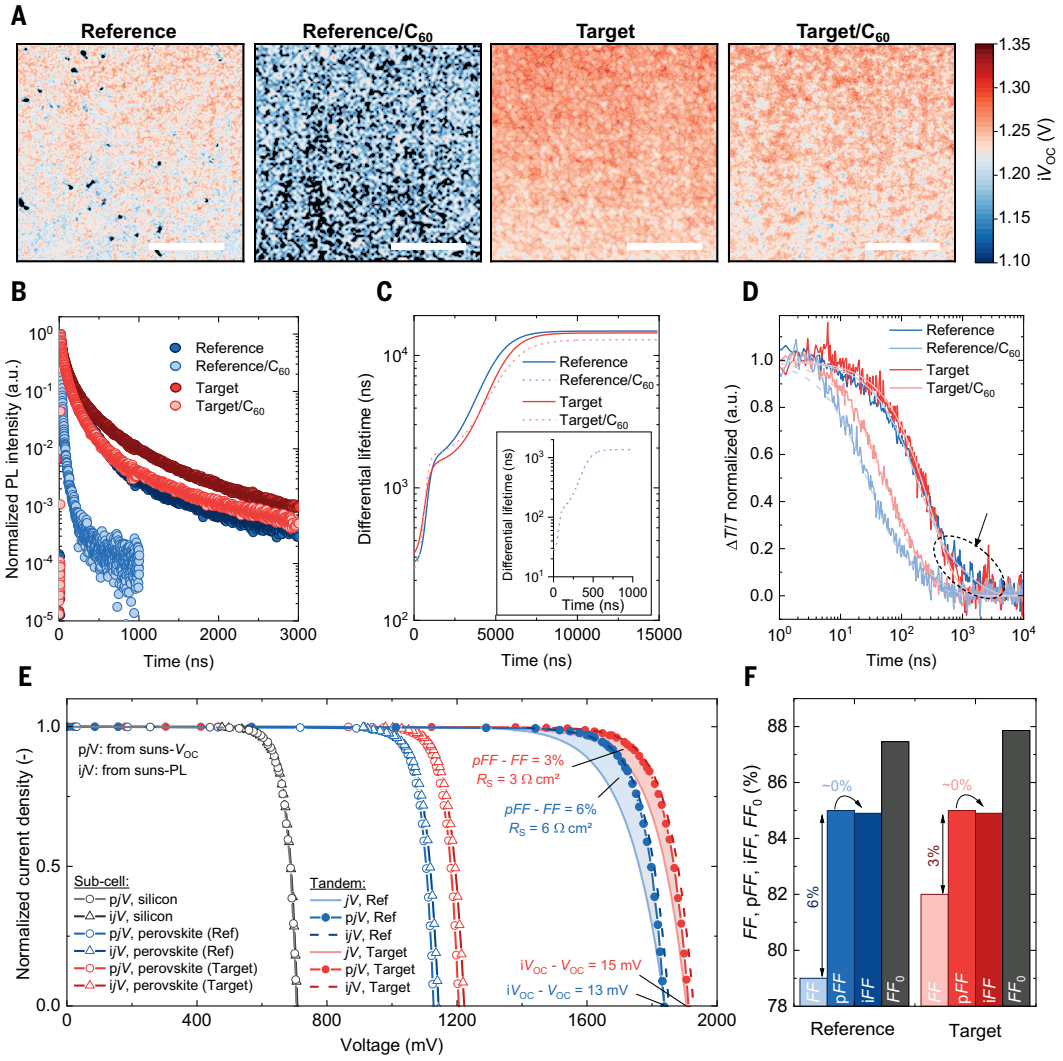


Fig. 3. Passivation and conductivity effects. (**A**) iV_{OC} images of reference and target films, from hyperspectral imaging. Scale bars, 100 μm. (**B** and **C**) Time-resolved PL decay (B) and differential lifetime (C) of reference and target films, from TRPL measurements. (**D**) Normalized bandgap bleach decay of reference and target films from transient absorption measurements. The dashed oval delineates the shallow trap signal in the reference stack. (**E**) *J-V*, pseudo-*J-V*, implied-*J-V* reverse scan curves of representative reference, and target fully textured perovskite/silicon tandem solar cells from *J-V*, suns-V_{OC}, and subcell-selective spatially resolved suns-PL imaging measurements. (**F**) Extracted FF from *J-V* measurements, pseudo-FF from suns-V_{OC}, implied-FF from suns-PL, and radiative FF (FF0) of representative reference and target fully textured perovskite/silicon tandem solar cells.

of carrier densities. By contrast, terahertz spectroscopy measurements of reference and PDAI-passivated perovskite films showed comparable mobility results in the range of 16.5 to 18.5 $cm^2 V^{-1} s^{-1}$ (fig. S15). This was consistent with values extracted from space-charge–limited current (SCLC) measurements (fig. S16) and confirmed the major contribution of enhanced electron concentration in the target perovskite film (interface and bulk) as the underlying factor behind the observed reduction in lumped series resistance.

Extending the optoelectrical study, we performed spatially resolved subcell-selective implied-FF (iFF) imaging through suns-PL characterization (*50*). Specifically, the illumination intensity was varied and the collected PL images were converted to iV_{OC} images (*51*), to then construct global suns-iV_{OC} curves for the tandem cells (Fig. 3E). The calculated iFF for reference and PDAI-passivated devices were similar (85%, Fig. 3F), indicating little impact of reduced interfacial recombination on the FF of the device. Furthermore, the difference between iV_{OC} and V_{OC} was negligible across both systems (~15 mV), indicating no change in selectivity losses (Fig. 3E). We concluded that the enhanced FF in the target device primarily resulted from reduced transport losses through increased electron concentration of the absorber, aligning with optoelectrical simulation. This factor is often overlooked compared to the commonly reported reduction in nonradiative recombination.

Performance of photovoltaic devices

A schematic of the tandem stack is shown in Fig. 4A, and a typical cross-sectional SEM image of the device is shown in Fig. 4B. Tandem solar cells with optimized PDAI concentration increased the PCE in the best devices from 29.1 to 32.3% (fig. S17), through improvements particularly in V_{OC} (1.83 to 1.94 V) and FF (79.4 to 81.6%), as supported by the statistical comparison in Fig. 4C (optimization steps in figs. S18 and S19). The short-circuit current density (J_{SC}) remained unchanged, which is consistent with external quantum efficiency (EQE) and reflectance results (fig. S17, B and C). The statistical comparison over five consecutive batches shows a reproducible 2.5%$_{abs}$ increase in *PCE* for PDAI-passivated devices over reference devices (fig. S20).

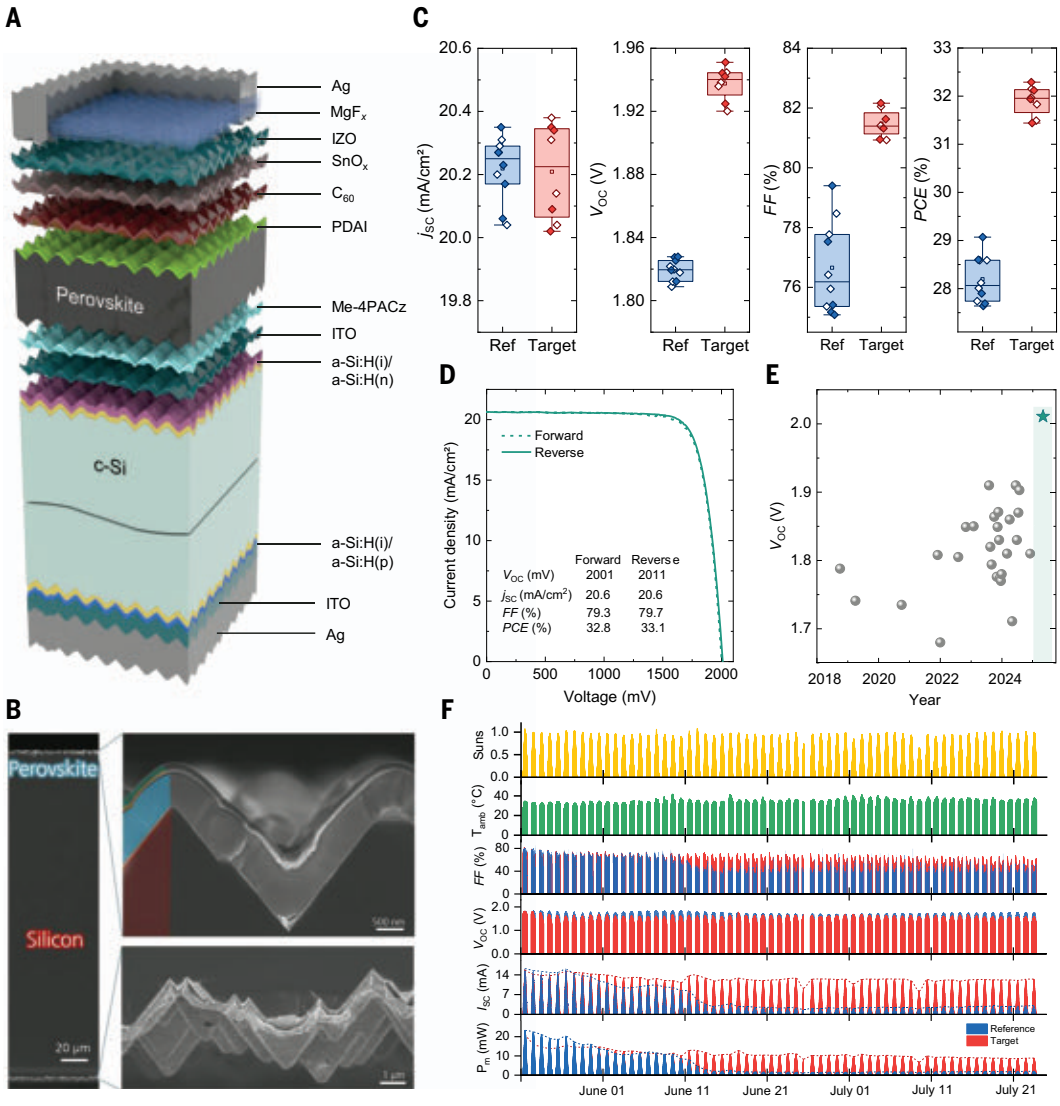


Fig. 4. Perovskite/silicon tandem solar cell performance and stability. (**A**) Sketch of the tandem solar cell structure. (**B**) Cross-sectional SEM image of the device. (**C**) PV parameters of reference and target devices (open symbols for forward scan and closed symbols for reverse scan). (**D**) *J-V* curves of the champion tandem solar cell with adapted bandgap [1.70 eV instead of 1.67 eV used in (C)]. (**E**) V_{OC} evolution of fully textured perovskite/silicon tandem solar cells. (**F**) MPP stability tracking of encapsulated devices outdoor (in the Red Sea region) with evolution of PV parameters, illumination intensity, and ambient temperature over the testing period (1500 hours).

The reproducibility of the strategy in different labs (located in different institutes, namely KAUST and Fraunhofer ISE) was shown to be consistent (fig. S21). One of the devices was measured by an accredited laboratory, obtaining a certified stabilized *PCE* of 31.6% (fig. S22). Among the target devices, we further increased the perovskite bandgap from 1.67 to 1.70 eV for enhanced current matching (*23*, *27*), leading to a champion device delivering a V_{OC} of 2.01 V and a *PCE* of 33.1% (Fig. 4D and fig. S23).

To evaluate the effect of passivation on the stability of (encapsulated) tandem cells (*52*), we conducted an outdoor test at the coast of the Red Sea. The test included all quadrants of the operational conditions such as continuous operation under illumination and heat (up to 40°C), and diurnal heat cycles. Reference and target devices (*J-V* curves after encapsulation in fig. S23) showed a comparable constant trend in V_{OC}, which hints toward no generation of new defects in the perovskite absorber or loss in selectivity (Fig. 4F). However, the daily peak J_{SC} gradually diverged over time approaching a null value in the reference device after nearly 500 hours. By contrast, a more stable current output was recorded for the PDAI-passivated device over the following 1000 hours. These results were reproducible across multiple tandem solar cells (fig. S25).

Previous studies have indicated that degradation of perovskite solar cells under operational conditions is mainly driven by mobile ion-induced internal field screening, which leads to a large reduction in the current and thereby power output (*53*–*55*). By performing bias-assisted charge extraction measurements (BACE), we found that the concentration of mobile ions after the stability test in the reference device (6.5×10^{17} cm^{-3}) was nearly three times that of the target device (2.6×10^{17} cm^{-3}), which may constitute a clear rational for the accelerated degradation of reference devices (fig. S26). The outdoor stability test was complemented with damp heat testing, whereby devices were aged at elevated temperature and relative humidity (85°C, 85%) for 1000 hours, corresponding to ISOS-D-3 (fig. S27). An improved resilience to FF and V_{OC} losses was found in the target devices, which minimized *PCE* degradation from 73.7 to 83.4% relative to initial performance, thereby substantiating the enhanced stability of the passivated fully textured tandem solar cell.

Discussion

We demonstrated that the effect of work function engineering strategies extends beyond the interface to the bulk of the perovskite absorber. This effect is well known from other photovoltaic technologies [e.g., field-effect passivation by negative charges in Al_2O_3 layers at boron-doped emitters (*56*) or base material (*57*) in modern silicon solar cells]. However, in these devices, the accumulation region only changes the properties of a narrow region (around 1% of the complete device) owing to the high doping concentration of the silicon absorber.

In our study, we find a completely different situation. Because of the intrinsic properties of the perovskite bulk, the induced electron accumulation region extended across the entire perovskite absorber, giving rise to a "deep" field-effect passivation (supplementary text), which in turn raised the absorber's conductivity and thus enhanced electron transport. Initially validated through optoelectrical simulations, this finding was further supported by subcell-selective, spatially resolved iFF imaging and suns-V_{OC} measurements, which collectively indicate that the improvement in FF is predominantly driven by the reduction of charge carrier transport losses—an effect that has so far been overlooked compared to the mitigation of interfacial recombination.

Moreover, as revealed by AM-KPFM and iV_{OC} analysis, although PDAI homogenizes the surface potential of the perovskite, leading to a slight reduction of 20 mV in chemical passivation losses, the primary increase in V_{OC} arose from the reduced conduction band offset (110 meV), minimizing interfacial recombination.

Notably, the interaction between PDAI and the surface of the hybrid route–deposited perovskite resulted in a dipole oriented in the opposite direction compared to that observed in solution-processed perovskites (*8*, *15*, *16*, *40*–*43*), emphasizing the role of the perovskite's surface termination on dipole formation. Despite the difference in dipole direction and thus the change in the perovskite work function, the conduction band offset was reduced in both cases. This result indicated that the critical metric for the success of such passivation strategies was increasing the ionization potential rather than controlling the dipole direction, for minimizing $\Delta E_{C,ETL}$. Furthermore, our study demonstrates that, independently of the work function value (i.e., slightly more or less *n*-type target perovskite), as long as the Fermi level of the perovskite absorber is positioned mid-gap (i.e., intrinsic material), the induced electron junction with reduced $\Delta E_{C,ETL}$ will lead to the accumulation of a high electron density throughout the perovskite absorber, improving FF and V_{OC}. This effect explains the consistent success of this passivation molecule across various perovskite compositions.

At this advanced stage of ETL optimization, the optoelectrical study indicated that further reducing the conduction band offset will compromise the HTL because of the high Debye length of the intrinsic perovskite bulk, highlighting the need for simultaneous engineering of both interfaces.

Implementing the described strategies in a fully textured perovskite/silicon tandem solar cell enabled us to achieve an exceptionally high V_{OC} of 2.01 V. The 100-mV improvement in V_{OC} compared to that of previous devices positions the industry-compatible architecture on par with its record-setting counterparts. Furthermore, it demonstrates that there is no inherent limitation for interfacial passivation caused by the increased perovskite surface area on industrially produced silicon solar cells with large pyramid heights (*25*, *26*).

Finally, the increased PCE provided by passivation did not compromise stability. Instead, the modified perovskite/C_{60} interface immobilized ion migration across the interface as quantified with BACE measurements, enhancing the device's operational stability. Despite these encouraging results, we believe that further studies focusing on the bulk quality of the perovskite material prepared via the hybrid method may offer promising prospects for future commercialization of the technology.

Materials and methods are available as supplementary materials.

REFERENCES AND NOTES

1. J. J. Cordell, M. Woodhouse, E. L. Warren, Technoeconomic analysis of perovskite/silicon tandem solar modules. *Joule* **9**, 101781 (2025). doi: 10.1016/j.joule.2024.10.013
2. E. Aydin *et al.*, Pathways toward commercial perovskite/silicon tandem photovoltaics. *Science* **383**, eadh3849 (2024). doi: 10.1126/science.adh3849; pmid: 38207044
3. E. Aydin *et al.*, Enhanced optoelectronic coupling for perovskite/silicon tandem solar cells. *Nature* **623**, 732–738 (2023). doi: 10.1038/s41586-023-06667-4; pmid: 37769785
4. P. Tockhorn *et al.*, Nano-optical designs for high-efficiency monolithic perovskite-silicon tandem solar cells. *Nat. Nanotechnol.* **17**, 1214–1221 (2022). doi: 10.1038/s41565-022-01228-8; pmid: 36280763
5. E. Ugur *et al.*, Enhanced cation interaction in perovskites for efficient tandem solar cells with silicon. *Science* **385**, 533–538 (2024). doi: 10.1126/science.adp1621; pmid: 39088622
6. D. Turkay *et al.*, Synergetic substrate and additive engineering for over 30%-efficient perovskite-Si tandem solar cells. *Joule* **8**, 1735–1753 (2024). doi: 10.1016/j.joule.2024.04.015
7. A. Al-Ashouri *et al.*, Monolithic perovskite/silicon tandem solar cell with >29% efficiency by enhanced hole extraction. *Science* **370**, 1300–1309 (2020). doi: 10.1126/science.abd4016; pmid: 33303611
8. J. Liu *et al.*, Perovskite/silicon tandem solar cells with bilayer interface passivation. *Nature* **635**, 596–603 (2024). doi: 10.1038/s41586-024-07997-7; pmid: 39236747
9. S. Mariotti *et al.*, Interface engineering for high-performance, triple-halide perovskite-silicon tandem solar cells. *Science* **381**, 63–69 (2023). doi: 10.1126/science.adf5872; pmid: 37410849
10. A. A. Said *et al.*, Sublimed C_{60} for efficient and repeatable perovskite-based solar cells. *Nat. Commun.* **15**, 708 (2024). doi: 10.1038/s41467-024-44974-0; pmid: 38267408
11. J. Liu *et al.*, Efficient and stable perovskite-silicon tandem solar cells through contact displacement by MgF_x. *Science* **377**, 302–306 (2022). doi: 10.1126/science.abn8910; pmid: 35737811
12. D. Menzel *et al.*, Field Effect Passivation in Perovskite Solar Cells by a LiF Interlayer. *Adv. Energy Mater.* **12**, 2201109 (2022). doi: 10.1002/aenm.202201109
13. R. R. Pradhan *et al.*, Elucidating the Role of Contact-Induced Gap States and Passivation Molecules at Perovskite/Metal Contacts. *ACS Appl. Energy Mater.* **6**, 4111–4118 (2023). doi: 10.1021/acsaem.3c00292
14. J. Warby *et al.*, Understanding Performance Limiting Interfacial Recombination in pin Perovskite Solar Cells. *Adv. Energy Mater.* **12**, 2103567 (2022). doi: 10.1002/aenm.202103567
15. H. Chen *et al.*, Regulating surface potential maximizes voltage in all-perovskite tandems. *Nature* **613**, 676–681 (2023). doi: 10.1038/s41586-022-05541-z; pmid: 36379225
16. C. Liu *et al.*, Bimolecularly passivated interface enables efficient and stable inverted perovskite solar cells. *Science* **382**, 810–815 (2023). doi: 10.1126/science.adk1633; pmid: 37972154
17. Y. Wang *et al.*, Homogenized contact in all-perovskite tandems using tailored 2D perovskite. *Nature* **635**, 867–873 (2024). doi: 10.1038/s41586-024-08158-6; pmid: 39401514
18. N. Tucher *et al.*, Energy yield analysis of textured perovskite silicon tandem solar cells and modules. *Opt. Express* **27**, A1419–A1430 (2019). doi: 10.1364/OE.27.0A1419; pmid: 31684495
19. F. Gota, R. Schmager, A. Farag, U. W. Paetzold, Energy yield modelling of textured perovskite/silicon tandem photovoltaics with thick perovskite top cells. *Opt. Express* **30**, 14172–14188 (2022). doi: 10.1364/OE.447069; pmid: 35473167
20. A. Callies, M. Hanser, J. C. Goldschmidt, B. Bläsi, O. Höhn, Structuring of perovskite-silicon tandem solar cells for reduced reflectance and thermalization losses. *Opt. Express* **31**, 19428–19442 (2023). doi: 10.1364/OE.491371; pmid: 37381358
21. Z. Liu *et al.*, Strained heterojunction enables high-performance, fully textured perovskite/silicon tandem solar cells. *Joule* **8**, 2834–2850 (2024). doi: 10.1016/j.joule.2024.06.015
22. F. Sahli *et al.*, Fully textured monolithic perovskite/silicon tandem solar cells with 25.2% power conversion efficiency. *Nat. Mater.* **17**, 820–826 (2018). doi: 10.1038/s41563-018-0115-4; pmid: 29891887
23. O. Er-raji *et al.*, Tailoring perovskite crystallization and interfacial passivation in efficient, fully textured perovskite silicon tandem solar cells. *Joule* **8**, 2811–2833 (2024). doi: 10.1016/j.joule.2024.06.018
24. X. Zheng *et al.*, Solvent engineering for scalable fabrication of perovskite/silicon tandem solar cells in air. *Nat. Commun.* **15**, 4907 (2024). doi: 10.1038/s41467-024-49351-5; pmid: 38851760
25. X. Y. Chin *et al.*, Interface passivation for 31.25%-efficient perovskite/silicon tandem solar cells. *Science* **381**, 59–63 (2023). doi: 10.1126/science.adg0091; pmid: 37410835
26. D. Khan, G. Qu, I. Muhammad, Z. Tang, Z.-X. Xu, Overcoming Two Key Challenges in Monolithic Perovskite-Silicon Tandem Solar Cell Development: Wide Bandgap and Textured Substrate—A Comprehensive Review. *Adv. Energy Mater.* **13**, 2302124 (2023). doi: 10.1002/aenm.202302124
27. O. Er-raji *et al.*, Loss Analysis of Fully-Textured Perovskite Silicon Tandem Solar Cells: Characterization Methods and Simulation Towards the Practical Efficiency Potential. *Sol. RRL* **7**, 2300659 (2023). doi: 10.1002/solr.202300659
28. Sentaurus TCAD, *Technology Computer Aided Design (TCAD)* (Synopsys, Zürich, Switzerland, release U-2022.12-SP1).

29. C. Messmer *et al.*, Toward more reliable measurement procedures of perovskite-silicon tandem solar cells: The role of transient device effects and measurement conditions. *Prog. Photovolt. Res. Appl.* **33**, 126–142 (2025). doi: 10.1002/pip.3782
30. F. Zu *et al.*, Charge Selective Contacts to Metal Halide Perovskites Studied with Photoelectron Spectroscopy: X-Ray, Ultraviolet, and Visible Light Induced Energy Level Realignment. *Adv. Mater. Interfaces* **10**, 2300413 (2023). doi: 10.1002/admi.202300413
31. X.-Y. He *et al.*, Interfacial Field-Effect Enabling High-Performance Perovskite Photovoltaics. *Small* **21**, e2410310 (2025). doi: 10.1002/smll.202410310; pmid: 39703151
32. F. Peña-Camargo *et al.*, Revealing the doping density in perovskite solar cells and its impact on device performance. *Appl. Phys. Rev.* **9**, 021409 (2022). doi: 10.1063/5.0085286
33. B. Das, I. Aguilera, U. Rau, T. Kirchartz, Effect of Doping, Photodoping, and Bandgap Variation on the Performance of Perovskite Solar Cells. *Adv. Opt. Mater.* **10**, 2101947 (2022). doi: 10.1002/adom.202101947
34. S. Teale, M. Degani, B. Chen, E. H. Sargent, G. Grancini, Molecular cation and low-dimensional perovskite surface passivation in perovskite solar cells. *Nat. Energy* **9**, 779–792 (2024). doi: 10.1038/s41560-024-01529-3
35. S. Olthof *et al.*, Ultralow doping in organic semiconductors: Evidence of trap filling. *Phys. Rev. Lett.* **109**, 176601 (2012). doi: 10.1103/PhysRevLett.109.176601; pmid: 23215211
36. L. Canil *et al.*, Tuning halide perovskite energy levels. *Energy Environ. Sci.* **14**, 1429–1438 (2021). doi: 10.1039/D0EE02216K
37. O. Er-raji *et al.*, Towards Efficient and Industrially-Compatible Fully-Textured Perovskite Silicon Tandem Solar Cells: Controlled Process Parameters for Reliable Perovskite Formation. 5. *Prog. Photovolt. Res. Appl.* **33**, 86–99 (2023). doi: 10.1002/pip.3770
38. M. R. Ahmadian-Yazdi, F. Zabihi, M. Habibi, M. Eslamian, Effects of Process Parameters on the Characteristics of Mixed-Halide Perovskite Solar Cells Fabricated by One-Step and Two-Step Sequential Coating. *Nanoscale Res. Lett.* **11**, 408 (2016). doi: 10.1186/s11671-016-1601-8; pmid: 27639581
39. M. Wang, Y. Feng, J. Bian, H. Liu, Y. Shi, A comparative study of one-step and two-step approaches for $MAPbI_3$ perovskite layer and its influence on the performance of mesoscopic perovskite solar cell. *Chem. Phys. Lett.* **692**, 44–49 (2018). doi: 10.1016/j.cplett.2017.12.012
40. S. Hu *et al.*, Optimized carrier extraction at interfaces for 23.6% efficient tin–lead perovskite solar cells. *Energy Environ. Sci.* **15**, 2096–2107 (2022). doi: 10.1039/D2EE00288D
41. T. Ye *et al.*, Molecular Synergistic Effect for High Efficiency Monolithic Perovskite/Perovskite/Silicon Triple-Junction Tandem Solar Cells. *Adv. Energy Mater.* **14**, 2402491 (2024). doi: 10.1002/aenm.202402491
42. J. Wang *et al.*, Unraveling the Molecular Size Effect on Surface Engineering of Perovskite Solar Cells. *Small Methods* **8**, e2400043 (2024). doi: 10.1002/smtd.202400043; pmid: 38462962
43. Z.-R. Lan *et al.*, Surface Passivation with Diaminopropane Dihydroiodide for p-i-n Perovskite Solar Cells with Over 25% Efficiency. *Adv. Funct. Mater.* **34**, 2312426 (2024). doi: 10.1002/adfm.202312426
44. J. Liang *et al.*, Origins and influences of metallic lead in perovskite solar cells. *Joule* **6**, 816–833 (2022). doi: 10.1016/j.joule.2022.03.005
45. B. Roose, K. Dey, Y.-H. Chiang, R. H. Friend, S. D. Stranks, Critical Assessment of the Use of Excess Lead Iodide in Lead Halide Perovskite Solar Cells. *J. Phys. Chem. Lett.* **11**, 6505–6512 (2020). doi: 10.1021/acs.jpclett.0c01820; pmid: 32693601
46. B. Shi *et al.*, Unraveling the Passivation Process of PbI_2 to Enhance the Efficiency of Planar Perovskite Solar Cells. *J. Phys. Chem. C Nanomater. Interfaces* **122**, 21269–21276 (2018). doi: 10.1021/acs.jpcc.8b08075
47. F. Ye *et al.*, Overcoming C_{60}-induced interfacial recombination in inverted perovskite solar cells by electron-transporting carborane. *Nat. Commun.* **13**, 7454 (2022). doi: 10.1038/s41467-022-34203-x; pmid: 36460635
48. B. Krogmeier, F. Staub, D. Grabowski, U. Rau, T. Kirchartz, Quantitative analysis of the transient photoluminescence of $CH_3NH_3PbI_3/PC_{61}BM$ heterojunctions by numerical simulations. *Sustain. Energy Fuels* **2**, 1027–1034 (2018). doi: 10.1039/C7SE00603A
49. A. Cuevas, R. A. Sinton, Prediction of the open-circuit voltage of solar cells from the steady-state photoconductance. *Prog. Photovolt. Res. Appl.* **5**, 79–90 (1997). doi: 10.1002/(SICI)1099-159X(199703/04)5:2<79::AID-PIP155>3.0.CO;2-J
50. O. Fischer *et al.*, Revealing Charge Carrier Transport and Selectivity Losses in Perovskite Silicon Tandem Solar Cells. *Matter* 102404 (2025). doi: 10.1016/j.matt.2025.102404
51. O. Fischer *et al.*, Versatile Implied Open-Circuit Voltage Imaging Method and its Application in Monolithic Tandem Solar Cells. *Prog. Photovolt. Res. Appl.* **33**, 40–53 (2023). doi: 10.1002/pip.3754
52. A. A. Said *et al.*, Perovskite/Silicon Tandem Photovoltaics: Long-Term Stability through Interface Engineering. *Energy Fuels* **39**, 10134–10149 (2025). doi: 10.1021/acs.energyfuels.5c00511
53. J. Thiesbrummel *et al.*, Ion-induced field screening as a dominant factor in perovskite solar cell operational stability. *Nat. Energy* **9**, 664–676 (2024). doi: 10.1038/s41560-024-01487-w
54. S. Shah *et al.*, Impact of Ion Migration on the Performance and Stability of Perovskite-Based Tandem Solar Cells. *Adv. Energy Mater.* **14**, 2400720 (2024). doi: 10.1002/aenm.202400720
55. J. Thiesbrummel *et al.*, Universal Current Losses in Perovskite Solar Cells Due to Mobile Ions. *Adv. Energy Mater.* **11**, 2101447 (2021). doi: 10.1002/aenm.202101447
56. J. Benick *et al.*, High efficiency n-type Si solar cells on Al2O3-passivated boron emitters. *Appl. Phys. Lett.* **92**, 253504 (2008). doi: 10.1063/1.2945287
57. J. Schmidt *et al.*, Surface passivation of high-efficiency silicon solar cells by atomic-layer-10 deposited Al 2 O 3. *Prog. Photovolt. Res. Appl.* **16**, 461–466 (2008). doi: 10.1002/pip.823

ACKNOWLEDGMENTS

We thank L. V. T. Merino, P. Dally, A. Razzaq, T. G. Allen, A. El Labban, A. ur Rehman, S. Sarwade, S. Shikin, and R. Guadalquiver for fruitful discussions. **Funding:** This work was funded by the Fraunhofer LIGHTHOUSE PROJECT MaNiTU, the German Federal Ministry for Economic Affairs and Climate Action under contract nos. 03EE1086A (PrEsto), 03EE1182A, and 03EE1182B (Perle), and the King Abdullah University of Science and Technology (KAUST) Research Funding Office under Award nos. ORA-CRG10-2021-4669 and ORA-CRG10-2021-4681. H.P. thanks the Finnish Foundation for Technology Promotion, decision 8462, for funding the work. C.E.P. acknowledges support from KAUST Global Fellowship Program under award no. ORA-2022-5002. E.A. acknowledges financial support from the European Research Council (ERC) under the European Union's Horizon Europe Research and Innovation Program (INPERSPACE, grant agreement no. 101077006). C.M. acknowledges support from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under SPP 2196: Perovskite Semiconductors: From Fundamental Properties to Devices, project no. 402726906. J.B. acknowledges support The Vector Stiftung for funding her research group. **Author contributions:** O.E. conceived the project idea and coordinated the work. O.E. fabricated film stacks for characterization results; carried out BB-XRD, GI-XRD, GIWAXS, SEM, and hyperspectral PL imaging; fabricated tandem solar cells; performed *J-V* and *EQE* measurements; and wrote the first draft of the manuscript. C.M. carried out the optoelectrical simulations, helped in understanding the solar cell results, and wrote part of the manuscript. R.R.P. carried out DFT calculations and wrote part of the manuscript. O.F. carried out the suns-V_{OC} and suns-PL measurements, evaluated the data, and helped understanding of the solar cell results. V.H. carried out the QFLS analysis and helped with hyperspectral PL measurements. S.K. and V.H. conducted XPS and UPS measurements and evaluated the data. M.M. carried out the operational stability measurements and evaluated the data. M.L. and J.F. carried out the BACE measurements with the help of C.B. and Y.G. for experiment setup and sample preparation and evaluated the data. O.E. and Y.G. optimized the scalable PDAI deposition process. J.P.J. performed terahertz measurements and evaluated the data. O.M. conducted TRPL measurements and evaluated the data. H.P. carried out TA and TR measurements and evaluated the data. C.E.P. helped develop methodology for time-resolved spectroscopic characterization (THz, TRPL, TA, and TR) measurements and analysis. A.P. carried out AM-KPFM measurements, evaluated the data, and made the cell schematic in Fig. 4A. B.V. carried out STEM measurements and evaluated the data. S.Z. and E.U. helped with perovskite subcell fabrication. M.B. provided silicon bottom solar cells and A.R.P. helped with last fabrication steps. M.B. provided important scientific insights. C.E.P., R.A., E.A., J.S., F.S., M.S., U.S., F.L., A.A.S., J.B., P.S.C.S., S.D.W., and S.G. supervised the overall project, and U.S., F.L., S.D.W., and S.G. secured the funding. All authors contributed to the manuscript revision process. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adx1745
Materials and Methods; Supplementary Text; Figs. S1 to S29; References (*58*–*89*)

Submitted 4 March 2025; resubmitted 24 March 2025; accepted 6 August 2025; published online 4 September 2025

10.1126/science.adx1745

RESEARCH ARTICLE SUMMARY

MOLECULAR BIOLOGY

FIGNL1 inhibits homologous recombination in BRCA2 deficient cells by dissociating RAD51 filaments

Raviprasad Kuthethur *et al.*

Full article and list of author affiliations: https://doi.org/10.1126/science.adt1210

INTRODUCTION: Homologous recombination (HR) is essential for DNA double-strand break (DSB) repair, and Breast Cancer Gene 2 (BRCA2) facilitates RAD51 nucleofilament formation crucial for strand exchange in DSB repair. Cells lacking BRCA2 show defects in RAD51 loading, leading to impaired HR and genomic instability and predisposing them to cell death, drug sensitivity, and tumorigenesis. RAD51 loading requires stringent regulation as excessive RAD51 on chromatin causes genome instability. This regulation occurs through RAD51 dissociation from DNA, mediated by helicases during HR, suggesting a balance between RAD51 loading and removal.

RATIONALE: Pathogenic BRCA2 mutations are viable chemotherapy targets through synthetic lethal strategies targeting HR deficiency. However, some BRCA2-deficient patients develop resistance through unknown HR restoration mechanisms. We hypothesized that HR restoration might result from inactivating mechanisms dissociating RAD51 from DSBs. We investigated the role of the AAA+ ATPase Fidgetin Like 1 (FIGNL1) in facilitating the dissociation of RAD51 and assessed its impact on the restoration of homologous recombination in the context of BRCA2 deficiency.

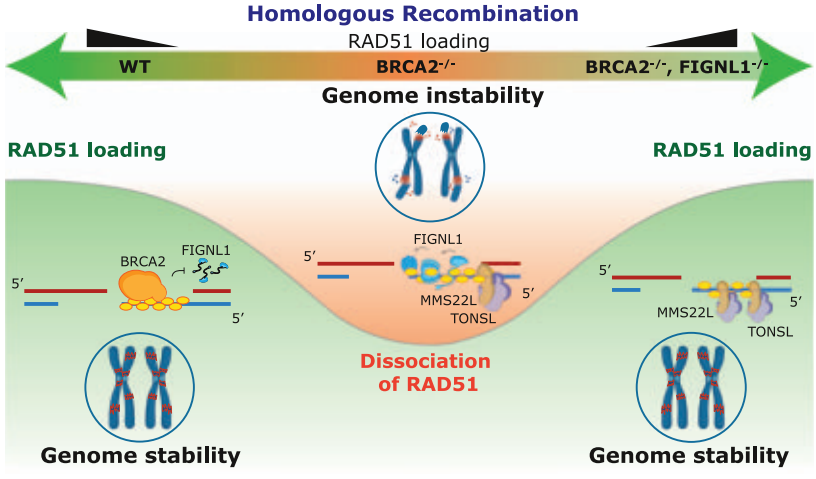
RESULTS: Our findings demonstrate that the loading of RAD51 at DSBs in BRCA2-deficient cells can be restored through the loss of the anti-recombinase FIGNL1. The genetic knockout of FIGNL1 in both mouse and human BRCA2-deficient cells reverses HR defects and mitigates genome instability, which correlates with increased chemoresistance to PARPi and platinum-based drugs. Additionally, the loss of FIGNL1 results in the synthetic viability of BRCA2-deficient mouse embryonic stem cells, with restored RAD51 loading at DSBs. These findings suggest that HR defects in BRCA2-deficient cells arise from the unregulated removal of RAD51 by FIGNL1, rather than from defective RAD51 loading.

We also show that FIGNL1 can dissociate RAD51 when assembled into a hexameric form, with initial interaction facilitated by FIGNL1's RAD51 binding domain (FRBD). The hexameric interactions between RAD51 and FIGNL1 are crucial for RAD51 removal from DSBs. We identified the MMS22L-TONSL complex as responsible for loading RAD51 in cells deficient in BRCA2 and FIGNL1. MMS22L-TONSL and RAD51 compete for the same binding site on FIGNL1's FRBD region. FIGNL1 loss results in increased MMS22L loading at DSBs, suggesting that FIGNL1 facilitates removal of RAD51 loaded by MMS22L-TONSL. The absence of MMS22L-TONSL in BRCA2/FIGNL1 double-deficient cells leads to HR defects, genome instability, and re-sensitization to chemotherapeutic agents.

CONCLUSION: Our findings define a pathway for modulating RAD51 levels during homologous recombination (HR), thereby refining the prevailing notion that HR defects in BRCA2-deficient cells are primarily attributable to RAD51 loading deficiencies. Our data suggest that the removal of RAD51 by FIGNL1 is responsible for HR defects in BRCA2-deficient cells. Furthermore, the identification of the MMS22L-TONSL complex as a potential mediator of RAD51 loading holds implications for understanding chemoresistance in BRCA2-deficient tumors and for developing strategies to re-sensitize these tumors to chemotherapy. □

*Corresponding author. Email: petr.cejka@irb.usi.ch (P.C.); a.raychaudhuri@erasmusmc.nl (A.R.C.) Cite this article as R. Kuthethur *et al.*, *Science* **390**, eadt1210 (2025). DOI: 10.1126/science.adt1210

Modulating regulation of RAD51 during homologous recombination. BRCA2 facilitates the loading and stabilization of RAD51 and suppresses the anti-recombinase function of FIGNL1. In cells lacking BRCA2, FIGNL1 dissociates RAD51 from double-strand breaks (DSBs), leading to a deficiency in homologous recombination and resulting in genomic instability. In BRCA2-deficient cells, the absence of FIGNL1 can restore RAD51 loading through the MMS22L-TONSL complex, thereby re-establishing homologous recombination and genomic stability.



MOLECULAR BIOLOGY

FIGNL1 inhibits homologous recombination in BRCA2 deficient cells by dissociating RAD51 filaments

Raviprasad Kuthethur[1,2], Safa Nasrin VZ[3]†, Satheesh Kumar Sengodan[4]†, Carmen Fonseca[1]†, Stefan Braunshier[3]†, Nupur Nagar[5]†, Ananya Acharya[3]†‡, Xingde Wang[1], Arjan F. Theil[1], Oluwakemi Ibini[1], Eleni-Maria Manolika[1], Kelly de Koning[1,2], Julien Dessapt[6,7]§, Amélie Fradet-Turcotte[6,7], Joyce H. G. Lebbink[1,2,8], Roland Kanaar[1,2], Krishna Mohan Poluri[5], Shyam K. Sharan[4], Petr Cejka[3]*, Arnab Ray Chaudhuri[1]*

Homologous recombination (HR) deficiency upon Breast Cancer Gene 2 (BRCA2) loss arises from defects in the formation of RAD51 nucleoprotein filaments. We demonstrate that loss of the anti-recombinase Fidgetin Like 1 (FIGNL1) retains RAD51 loading at DNA double-stranded breaks (DSBs) in BRCA2-deficient cells, leading to genome stability, HR proficiency, and viability of BRCA2-deficient mouse embryonic stem cells. Mechanistically, we demonstrate that strand invasion and subsequent HR defects upon BRCA2 loss primarily arise from the unrestricted removal of RAD51 from DSB sites by FIGNL1, rather than from defective RAD51 loading. Furthermore, we identify that the MMS22L-TONSL complex interacts with FIGNL1 and is critical for HR in BRCA2/FIGNL1 double-deficient cells. These findings identify a pathway for tightly regulating RAD51 activity to promote efficient HR, offering insights into mechanisms of chemoresistance in BRCA2-deficient tumors.

The homologous recombination (HR) pathway is essential for accurate repair of DNA double-stranded breaks (DSBs) (*1*). BRCA2 is essential for the HR process and interacts with RAD51, mediating nucleation of RAD51 nucleoprotein filaments on single-stranded DNA (ssDNA) by displacing RPA (*2*–*4*). This step is crucial for the strand invasion and DNA synthesis steps in HR. BRCA2 loss restricts RAD51 accumulation at DSBs, disrupting HR and leading to genome instability (*5*–*11*). Tumor cells with BRCA2 mutations show RAD51 loading and HR deficiencies (*12*–*15*), which are targets for chemotherapeutics such as poly(adenosine-diphosphate-ribose) polymerase inhibitors (PARPi) and DNA crosslinking agents (*15*–*17*). Although these treatments have shown success, some tumors develop resistance through compensatory mechanisms (*13*, *18*, *19*). No compensatory mechanisms of RAD51 loading and HR have been identified upon BRCA2 loss, highlighting its crucial role in HR.

The MMS22L-TONSL heterodimeric complex regulates HR by controlling RAD51 dynamics at replication-induced DSBs (*20*–*22*). This heterodimer interacts with RPA and contains a RAD51 interaction motif, facilitating RAD51-mediated strand invasion (*22*). TONSL interacts with ASF1 and CAF1 and recognizes H4K20me0 at DSBs, enabling MMS22L-mediated RAD51 loading on nascent DNA at DSBs (*21*, *23*, *24*). Loss of MMS22L-TONSL does not affect BRCA2 localization to DSBs, suggesting that it either functions downstream or separately in RAD51-mediated HR (*21*). Nevertheless, the relationship between BRCA2 and MMS22L-TONSL-dependent pathways in HR remains poorly understood.

The regulation of RAD51 loading and stabilization is critical as excessive accumulation of RAD51 on chromatin is associated with genome instability and cell death (*25*, *26*). This regulation is achieved through dissociation of RAD51 from DNA or by disruption of recombination intermediates through various helicases and translocases, including the RECQ family helicases, PARI, RTEL1, and FBH1 (*27*–*32*). FIGNL1 (fidgetin-like 1), a member of the AAA+ ATPase family, plays a major role in maintaining genome stability (*33*). Depletion of FIGNL1 results in accumulation of RAD51 and genome instability characterized by persistent ultrafine chromosome bridges (*33*). The interaction between FIGNL1 and FIRRM/FLIP is crucial for DNA crosslink repair and homologous recombination (*34*–*39*). FIGNL1 is also crucial for meiotic recombination, as its absence results in impaired spermatogenesis due to RAD51 hyperaccumulation (*40*, *41*). Mutations in the *Arabidopsis* BRCA2 paralog can also suppress the meiotic phenotypes associated with loss of FIGL1, underscoring the importance of FIGNL1 in HR (*42*). Nevertheless, the underlying mechanisms governing RAD51 loading and dissociation, as well as their effects on HR fidelity and genomic stability, remain to be elucidated.

Results

FIGNL1 loss restores RAD51 loading at DSBs upon BRCA2 deficiency

Loss of FIGNL1 increases chromatin association of RAD51 and is not compatible with cellular viability (*40*, *43*). Therefore, we genetically inactivated FIGNL1 using CRISPR-Cas9 in RPE1 hTERT $p53^{-/-}$ cells with doxycycline (Dox)-inducible shRNA against BRCA2 (wild type (WT) or shBRCA2) (*44*) (fig. S1, A and B). FIGNL1 inactivation was also performed in the *Brca2*$^{-/-}$*; Trp53*$^{-/-}$ mouse mammary tumor cell line KB2P1.21 (*Brca2*$^{-/-}$) and the KB2P1.21 reconstituted with wild type *Brca2* (WT/KB2PR) derived from *K14cre; Trp53*$^{F/F}$*;Brca2*$^{F/F}$ mouse tumor model (*45*) (fig. S1, C and D). Inactivation of FIGNL1 in a p53-deficient background yielded viable cells without substantial changes in cell cycle profiles, indicating that genome instability from FIGNL1 loss is tolerated upon p53 deficiency (fig. S1, E and F). We then assessed RAD51 accumulation with or without DSBs upon FIGNL1 loss in BRCA2-proficient or deficient conditions. Upon treatment with ionizing radiation (IR) to induce DSBs, we analyzed chromatin-bound RAD51 foci using quantitative immunofluorescence-based cytometry (QIBC) (*46*, *47*). FIGNL1 loss in both cell lines resulted in significantly higher RAD51 accumulation on chromatin, apparent under untreated conditions and enhanced by IR, correlating with increased H2AX colocalization (Fig. 1A and fig. S2, A and B). BRCA2 loss resulted in defective RAD51 accumulation as expected (Fig. 1A and fig. S2, A and B). Loss of FIGNL1 in BRCA2-deficient cells considerably rescued the RAD51 focus formation defect upon IR induced DSB induction (Fig. 1A, fig. S2, A and B). Next, to test whether FIGNL1 loss increases RAD51 focus formation upon site-specific DSBs, we expressed the AsiSI restriction enzyme in mouse mammary tumor lines, in which DSBs can be induced by Dox and 4-hydroxytamoxifen (4-OHT) treatment (Fig. 1B and fig. S2C) (*48*). AsiSI induction resulted

[1]Department of Molecular Genetics, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, the Netherlands. [2]Oncode Institute, Erasmus University Medical Center, Rotterdam, the Netherlands. [3]Institute for Research in Biomedicine, Università della Svizzera italiana (USI), Faculty of Biomedical Sciences, Bellinzona, Switzerland. [4]Mouse Cancer Genetics Program, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD, USA. [5]Department of Biosciences and Bioengineering, Indian Institute of Technology Roorkee, Uttarakhand, India. [6]Department of Molecular Biology, Medical Biochemistry and Pathology, Université Laval, Québec, Canada. [7]CHU de Québec Research Center-Université Laval, Laval University Cancer Research Center, Québec, Canada. [8]Department of Radiotherapy, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, the Netherlands.
***Corresponding author. Email: petr.cejka@irb.usi.ch (P.C.); a.raychaudhuri@erasmusmc.nl (A.R.C.)** †These authors contributed equally to this work. ‡Present address: Division of Radiation and Genome Stability, Department of Radiation Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA. §Present address: Centre for Molecular and Systems Biology, Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON, Canada.

in similar levels of DSBs across the tested cell lines, measured by 53BP1 foci using QIBC (fig. S2C). Site-specific DSBs showed substantially higher RAD51 foci upon FIGNL1 loss compared with WT cells. BRCA2 abrogated RAD51 focus formation, which was restored to WT levels upon concomitant FIGNL1 loss (Fig. 1B). Furthermore, we also confirmed this effect in U2OS-DIvA cells by downregulating BRCA2, FIGNL1, or both with 4-OHT treatment (*49*). FIGNL1 loss in BRCA2-deficient U2OS cells rescued RAD51 focus formation defects (fig. S2D). These data suggest that RAD51 accumulation defects from BRCA2 loss could at least in part result from RAD51 removal by FIGNL1, and not defective RAD51 loading.

FIGNL1 interacts with FIRRM/FLIP and disassembles RAD51 nucleoprotein filaments (*35*–*39*). To test whether DNA binding of FIGNL1 is required, we purified an N-terminal–truncated fragment of FIGNL1 (FIGNL1-dNLS) lacking 1 to 120 residues, which disrupts FIRRM/FLIP binding but remains functional (fig. S3A) (*35*). FIGNL1-dNLS showed weak binding to both ssDNA and double-stranded DNA (dsDNA), with binding observed only at high concentrations (fig. S3, B and C). Although saturating RPA did not recruit FIGNL1 to ssDNA, at low RPA concentrations FIGNL1 enhanced RPA binding to ssDNA but not mitochondrial SSB (fig. S3, D and E). These data suggest that FIGNL1 disrupts RAD51 nucleofilaments through direct interaction with the RAD51 protein. FIGNL1 interacts with RAD51 (*38*). In line, our in vitro assays using recombinant FIGNL1-dNLS and RAD51 and pulldown assays with reconstituted FIGNL1 from cells demonstrated robust FIGNL1-RAD51 interaction (fig. S3, F to H).

Next, to assess whether FIGNL1 disassembles RAD51 from ssDNA, we performed nuclease protection assays. Without RAD51, DNA2 nuclease considerably degraded DNA. The presence of RAD51 protected DNA, indicative of RAD51 binding. Addition of FIGNL1-dNLS to RAD51-bound ssDNA resulted in decreased protection of ssDNA in a concentration-dependent manner (fig. S4A), indicative of RAD51 dissociation from ssDNA. We then tested whether the dissociation of RAD51 from ssDNA by FIGNL1-dNLS also affected its DNA strand exchange activity. RAD51 catalyzes the exchange of a 3′-tailed oligonucleotide-based substrate with a radiolabeled dsDNA donor (Fig. 1C) (*5*). Increasing FIGNL1-dNLS concentrations inhibited the formation of the strand exchange product, indicating that RAD51 filament disruption results in DNA strand exchange defects (Fig. 1C and fig. S4B). Furthermore, in D-loop assays (*50*), FIGNL1-dNLS inhibited D-loop formation catalyzed by RAD51 and RAD54 (fig. S4, C and D). The addition of AMP-PNP to the strand exchange reaction ablated the ability of FIGNL1-dNLS to inhibit the RAD51-mediated strand exchange reaction, indicating the critical role of the FIGNL1-dNLS ATPase domain in dissociating RAD51 (fig. S4E). Next, we tested whether ATP hydrolysis by RAD51 was required for its removal by FIGNL1-dNLS. We employed RAD51-K133R, a mutant defective in ATP hydrolysis (*51*, *52*). FIGNL1-dNLS inhibited DNA strand exchange to an extent similar to that of wild-type RAD51, indicating that FIGNL1-dNLS does not dissociate RAD51 by stimulating its ATP hydrolytic activity (Fig. 1D and fig. S4F). Our immunofluorescence data suggest that BRCA2 could prevent FIGNL1-mediated RAD51 dissociation (Fig. 1, A and B, and fig. S2, A to D). We next reconstituted DNA strand exchange with RPA and mini-BRCA2 (miBRCA2), containing the BRC1 and BRC2 motifs, DNA-binding, and C-terminal domains of BRCA2 (*53*) (fig. S5, A and B). miBRCA2 stimulated strand exchange in the presence of RAD51 and/or RPA (Fig. 1E and fig. S5C) (*5*, *6*, *54*, *55*). The presence of BRCA2 in the reaction prevented the inhibition of DNA strand exchange by FIGNL1-dNLS compared with RAD51 alone (Fig. 1E and fig. S5C), suggesting that BRCA2 counteracts FIGNL1-mediated RAD51 dissociation for efficient HR.

Finally, we tested whether the loss of FIGNL1 restores functional HR in BRCA2-deficient cells. We performed siRNA-mediated downregulation of BRCA2, FIGNL1, and BRCA2/FIGNL1 in U2OS DR-GFP cells (*56*). I-SceI endonuclease expression in FIGNL1-deficient cells showed a slight but notable increase in HR efficiency (Fig. 1F). BRCA2 down-regulation decreased HR efficiency as expected. However, FIGNL1 deficiency significantly rescued HR defects observed in BRCA2-deficient cells. Our data suggest that the HR defects upon BRCA2 loss result from unrestricted dissociation of RAD51 from DNA by FIGNL1 rather than through RAD51 nucleofilament loading and stabilization defects.

Loss of FIGNL1 rescues lethality of BRCA2 deficient ESC and resistance to chemotherapeutic drugs

Genetic knockout of either *Brca2* or *Fignl1* in mouse embryonic stem cells (mESCs) are incompatible with survival, possibly due to defects in the tight control of HR (*9*, *40*, *43*, *57*). Because loss of FIGNL1 in BRCA2-deficient cells rescues HR (Fig. 1F), we tested whether FIGNL1 deficiency rescues the lethality of BRCA2 deficient mESCs. We generated FIGNL1-deficient PL2F7 mESCs with one null and one conditional allele of *Brca2* (*Brca2*$^{f/-}$) (*58*) by targeting with short hairpin RNAs (shRNAs) and selecting colonies with residual levels of FIGNL1 (Fig. 2A and fig. S6A). Upon transfection of *Brca2*$^{f/-}$ mESCs with Cre-recombinase and HAT selection, no *Brca2*-null colonies were obtained, confirming BRCA2's essential role (Fig. 2B). However, targeting FIGNL1 with two separate shRNAs yielded 37.1 and 53.2% HAT-resistant *Brca2*-null colonies (Fig. 2B and fig. S6B). Additionally, IR-induced RAD51 focus formation was substantially rescued in *Brca2/Fignl1* double-deficient mESCs when compared with HR-defective *Brca2* hypomorphic mutant ESCs (R2336H) (*59*) (Fig. 2C), suggesting that the double deficient cells were competent for HR.

Because the loss of FIGNL1 upon BRCA2 deficiency resulted in the restoration of HR (Fig. 1F), we next assessed the levels of genome instability by analyzing metaphase spreads after PARPi (olaparib) and cisplatin treatment in mouse mammary tumors and RPE1 cells. BRCA2/FIGNL1 double deficiency significantly reduced chromosomal aberrations compared to BRCA2 loss alone (Fig. 2, D to E, and fig. S7A). Similar data were also observed in mESCs, which displayed restoration of genome stability upon PARPi treatment in *Brca2*$^{-/-}$ *shFignl1* double-deficient cells compared to *Brca2*^{R2336H} hypomorphic cells (fig. S7B). Finally, clonogenic survival assays in mouse mammary tumors and mESCs showed that FIGNL1 loss in BRCA2-deficient cells rescued cellular sensitivity to PARPi and platinum drugs (Fig. 2, F to G, and fig. S7, C to F). These data suggest that restored HR in BRCA2-deficient cells upon FIGNL1 loss restores RAD51 loading, driving synthetic viability and chemoresistance.

FIGNL1 interaction with RAD51 is critical for its regulation

To gain further insights into the molecular interactions involving FIGNL1, we performed mass photometry experiments with FIGNL1-dNLS, which revealed the existence of various forms including monomer and hexameric species of FIGNL1 (fig. S8, A and B) consistent with recent reports (*35*, *43*). AlphaFold3-predicted hexameric structure of full-length FIGNL1 revealed that the N-terminal region is likely to be unstructured as indicated by predicted local distance difference test (pLDDT) < 50 and that oligomerization likely occurs between the C-terminal domains (pLDDT > 70) in the presence of Mg^{2+} and ATP (fig. S8, C and D, and table S1). Our mass photometry and pulldown experiments revealed predominantly a monomer-monomer RAD51-FIGNL1 interaction and possibly a FIGNL1 hexamer-RAD51 monomer interaction (fig. S9, A to D), aligning with recent evidence (*35*, *43*). The data indicate that the interaction between the RAD51 monomer and the FIGNL1 hexamer may be transient. We further modeled the interaction between FIGNL1 hexamer and RAD51 using AlphaFold3. The predicted structure indicated that RAD51 is embedded into the hexameric pore of FIGNL1 (a hallmark of many AAA+ ATPases) through the N-terminal end of RAD51 (fig. S10, A and B, and table S2). The model revealed three FIGNL1 regions, including FIGNL1-RAD51 binding domain (FRBD), Pore loop 1 and Pore loop 2 contributing to the

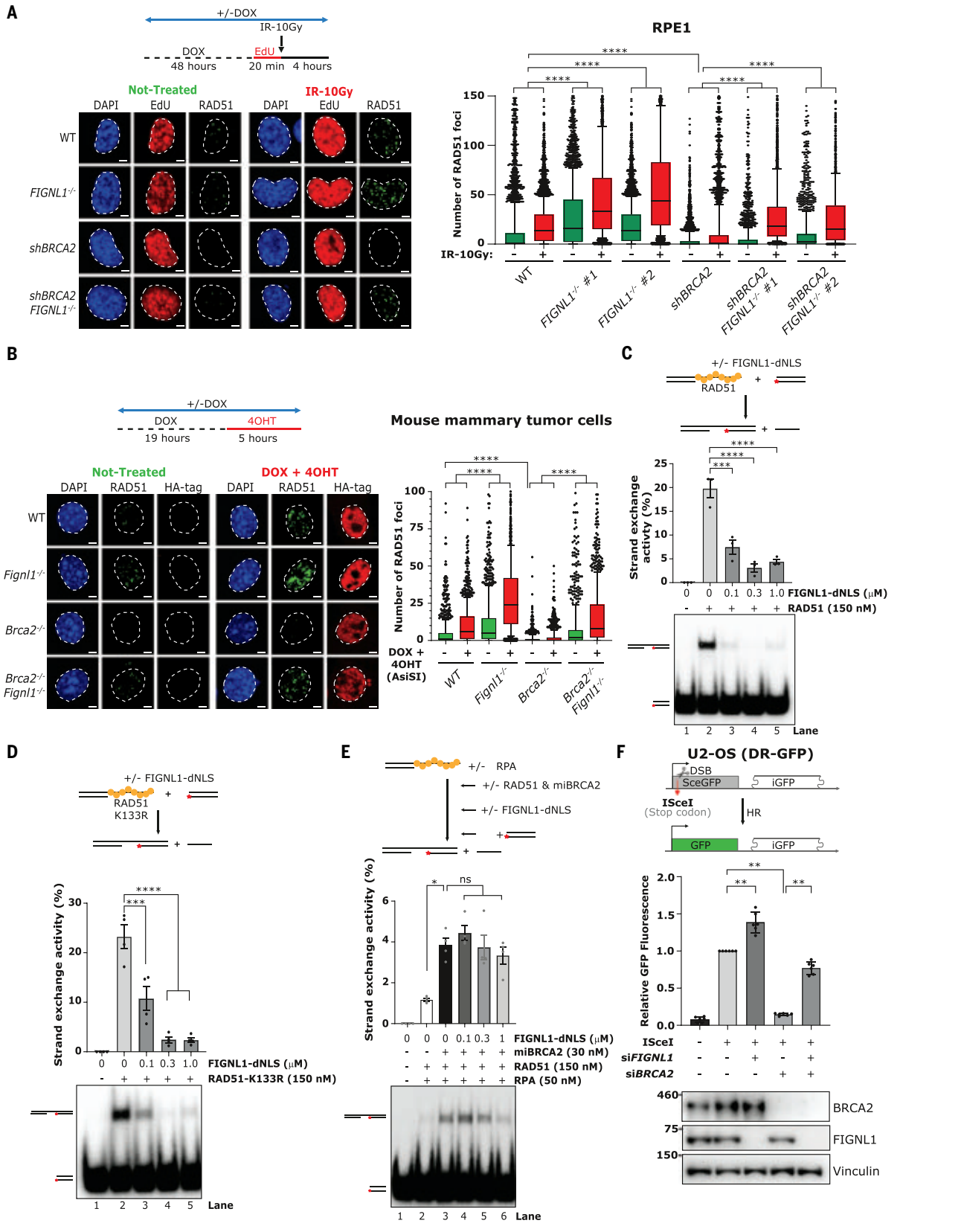


Fig. 1. Loss of FIGNL1 in BRCA2 deficient cells rescues RAD51 mediated HR. (**A**) (Top) Schematics showing the experimental setup for the high-content, automated RAD51 immunofluorescence (IF) in human RPE1 cells. 2 μg/ml of doxycycline for 48 hours to induce the knock-down of BRCA2, followed by EdU labelling and induction of DSB via 10 Gy of x-ray treatment. (Left) Representative IF images showing EdU and RAD51 foci in not-treated and ionizing radiation treated cells; scale bar, 5 μm. (Right) Automated Quantitative Image Based Cytometry (QIBC) plot of the RAD51 foci per nucleus; scatter box plot showing the foci distribution between the 10 to 90th percentiles from 2000 cells (experiment was repeated three times). *P*-values from Mann-Whitney test (two-tailed) comparing between respective not-treated or IR-treated groups (****: <0.0001).

(**B**) (Top) Schematics showing the experimental setup for RAD51-IF in mouse mammary tumor cells expressing AsiSI with HA-tag. 24 hours of doxycycline treatment to induce AsiSI-HA expression with last 5 hours to translocate it to the nucleus by 400 nM 4-hydroxytamoxifen (4OHT). (Left) Representative IF images of the RAD51 and HA-tag in not-treated and DOX+4OHT cells; scale bar, 5 μm. (Right) Automated QIBC plots of the RAD51 foci per nucleus; scatter box plot showing the foci distribution between the 10 to 90th percentiles from 1500 cells (experiment was repeated three times). *P*-values from Mann-Whitney test (two-tailed) comparing between respective not-treated or DOX+4OHT treated groups (****: <0.0001). (**C**) (Top) Schematics of the DNA strand exchange assays. The assays were performed by adding 150 nM purified RAD51 to the DNA substrates in the presence of increasing concentrations of purified FIGNL1-dNLS. The red asterisk indicates the position of the radioactive label. (Bottom) Quantification of the absolute DNA strand exchange activity showing the band intensity of the exchanged DNA; error bars are indicative of mean ± SEM from three independent experiments ($n = 3$). *P*-values from Dunnett's multiple comparison one-way ANOVA comparing between indicated groups (***: <0.001; ****: <0.0001). (**D**) (Top) Schematics of the DNA strand exchange assays. The assays were performed by adding 150 nM purified RAD51-K133R to the DNA substrates in the presence of increasing concentrations of purified FIGNL1-dNLS. The red asterisk indicates the position of the radioactive label. (Bottom) Quantification of the absolute DNA strand exchange activity showing the band intensity of the exchanged DNA; error bars are indicative of mean ± SEM from four independent experiments ($n = 4$). *P*-values from Dunnett's multiple comparison one-way ANOVA comparing between indicated groups (***: <0.001; ****: <0.0001). (**E**) (Top) Schematics of the DNA strand exchange assays. The assays were performed by adding 150 nM RAD51 and 30 nM miBRCA2 to the DNA substrates pre-incubated with 50 nM RPA. The reactions were then supplemented with the increasing concentrations of purified FIGNL1-dNLS and labelled donor oligos. The red asterisk indicates the position of the radioactive label. (Bottom) Quantification of the absolute DNA strand exchange activity showing the band intensity of the exchanged DNA; error bars are indicative of mean ± SEM from 4 independent experiments ($n = 4$). *P*-values from Dunnett's multiple comparison one-way ANOVA comparing between indicated groups (*: <0.05). (**F**) (Top) Schematics of DR-GFP assay showing the strategy behind the experiment. Mutant Sce-GFP containing ISceI recognition site and stop codon is repaired by the downstream iGFP to restore the GFP fluorescence in the cells. (Center) U2-OS cells with a stable integrated DR-GFP system were transfected for 72 hours with the indicated siRNAs and underwent 48 hours of ISceI treatment to measure the levels of GFP fluorescence relative to the WT cells undergoing ISceI treatment. Error bars are representative of the mean ± SD from three independent experiments ($n = 3$). *P*-values from unpaired *t*-test comparing between indicated groups (**: <0.01). (Bottom) Western blot analysis showing the levels of FIGNL1 and BRCA2 from the corresponding cells used for FACS analysis.

FIGNL1-RAD51 complex. Both Pore loops interact with RAD51's N-terminal region (table S2), as reported recently (*43*). The FRBD region was predicted to interact with two loops between Walker A and Walker B of the RAD51 ATPase domain with low confidence (pLDDT < 50) owing to the N-terminal region of FIGNL1 being unstructured. All five models generated, however, consistently predicted the interaction of RAD51 to the FRBD and the Pore loop regions with favorable binding energy (fig. S10, A and B, and table S2). Additionally, since the FRBD domain appears to be the primary RAD51 binding region on FIGNL1 (*38*), we used AlphaFold2 multimer to model RAD51 binding to the FRBD region. AlphaFold2 multimer prediction revealed a FIGNL1 FRBD-RAD51 dimer complex in which FRBD interacted with both RAD51 monomers (fig. S11, A and B, and table S3). These predictions suggest that although FRBD domain of FIGNL1 can bind to a RAD51 dimer, only one RAD51 molecule can interact with the pore loops of FIGNL1 in its hexameric form.

To validate AlphaFold3 predictions of the FIGNL1-RAD51 interactions, we generated mutants targeting conserved residues at interfaces. For in vitro reactions, mutants were generated in the FIGNL1-dNLS construct including dFRBD (lacking residues 295 to 344), FxxA mutation (F295E), Pore loop 1+2 mutations (K473E+W474A+H514A), dPore loop (lacking residues 461 to 562) predicted to affect RAD51 interaction (fig. S12, A to I). To evaluate the function of the catalytic domain of the FIGNL1- and RAD51-interacting region, we engineered several mutants by combining segments of the ATPase and RAD51 interaction domains. These include ATPase mutations (K447A+D500A), ATPase + Pore loop 1+2 mutations, F295E + ATPase + Pore loop 1+2 mutations, and dFRBD + ATPase + Pore loop 1+2 mutations (fig. S12, A to I). Testing these mutants revealed that FRBD deletion in FIGNL1-dNLS abolished RAD51 binding, whereas the other mutants retained the interaction with RAD51 in vitro under physiological salt conditions (Fig. 3A). This suggests multiple points of interaction between RAD51 and FIGNL1-dNLS, consistent with AlphaFold3 predictions. Notably, the ATPase and dPore loop mutants showed increased DNA binding compared to FIGNL1-dNLS (fig. S12J), though the implications remain unclear. Our data demonstrate that RAD51 interacts through multiple contacts with FIGNL1-dNLS, primarily in its FRBD domain.

To assess the modeled interactions in cells, we reconstituted eGFP-tagged full-length (FL) FIGNL1, dFRBD, FxxA mutant, dNLS, ATPase, Pore loop 1+2 mutations, ATPase + Pore loop 1+2 mutations, and dPore loop in *FIGNL1*$^{-/-}$ RPE1 cells (fig. S13, A and B) and conducted immunoprecipitation (IP) experiments. All reconstituted FIGNL1 mutants exhibited reduced RAD51 interaction compared with full-length FIGNL1. The dNLS mutation, which disrupts the interaction of FIGNL1 with FIRRM, demonstrated the lowest reduction whereas the dFRBD and F295E mutations showed the highest reduction among the mutants (Fig. 3B). This confirmed FRBD as the most critical domain for RAD51 binding (*38*), though FIGNL1's interaction with partners like FIRRM remains essential for optimal RAD51 interaction within protein complexes.

Subsequently, we evaluated the FIGNL1-RAD51 interaction mutants (fig. S12, A to I) in disassembling RAD51 from ssDNA through nuclease protection assays. As previously demonstrated (fig. S4A), FIGNL1-dNLS led to reduced protection of ssDNA through RAD51 filament disassembly (Fig. 3C). All generated RAD51 interaction mutants exhibited significant impairment in the removal of RAD51 from ssDNA compared to FIGNL1-dNLS (Fig. 3C). The ssDNA deprotection impairment by FIGNL1-dNLS mutants was associated with reduced ATP hydrolysis activity in mutants targeting conserved ATPase motifs (Fig. 3D). These findings indicate that the AlphaFold3 predicted interfaces between hexameric FIGNL1 and RAD51 are essential for FIGNL1 function in the dissociation of RAD51 filaments from DNA.

FIGNL1 association with RAD51 is essential for its removal and restricting strand exchange

Based on our in vitro biochemical data and cellular reconstitution experiments of FIGNL1-RAD51 interface mutants, we assessed the impact of these mutants in RAD51 focus accumulation after IR treatments. The reconstitution of FIGNL1-FL led to *FIGNL1*$^{-/-}$ and shBRCA2 *FIGNL1*$^{-/-}$ cells exhibiting reduced levels of RAD51 foci, akin to the conditions observed in WT or BRCA2-deficient cells, respectively. (Fig. 4, A and B). Reconstitution with all the mutants tested (fig. S13, A and B) failed to reduce the number of RAD51 foci in *FIGNL1*$^{-/-}$ or shBRCA2 *FIGNL1*$^{-/-}$ cells (Fig. 4, A and B). To test whether the mutants' failure to reduce RAD51 focus levels correlated with genome stability in shBRCA2 *FIGNL1*$^{-/-}$ cells, we measured chromosomal aberrations upon PARPi or cisplatin treatment. Reconstitution with FIGNL1-FL and treatment with either PARPi or cisplatin increased chromosomal aberrations in the BRCA2/FIGNL1 double-deficient cells (fig. S13C), in contrast to the FIGNL1 mutants (fig. S13C). Finally, we evaluated whether the RAD51 interaction mutants of FIGNL1-dNLS could impede RAD51-mediated DNA strand exchange products in our biochemical assay. Consistent with ssDNA protection data (Fig. 3C), each mutant showed impairment in disrupting DNA strand exchange reaction catalyzed by RAD51 (Fig. 4C). These data indicate that stabilization of FIGNL1 by FIRRM/FLIP and the interaction of RAD51 with FRBD, ATPase and Pore loop regions of FIGNL1 are critical for RAD51 dissociation from resected DSBs.

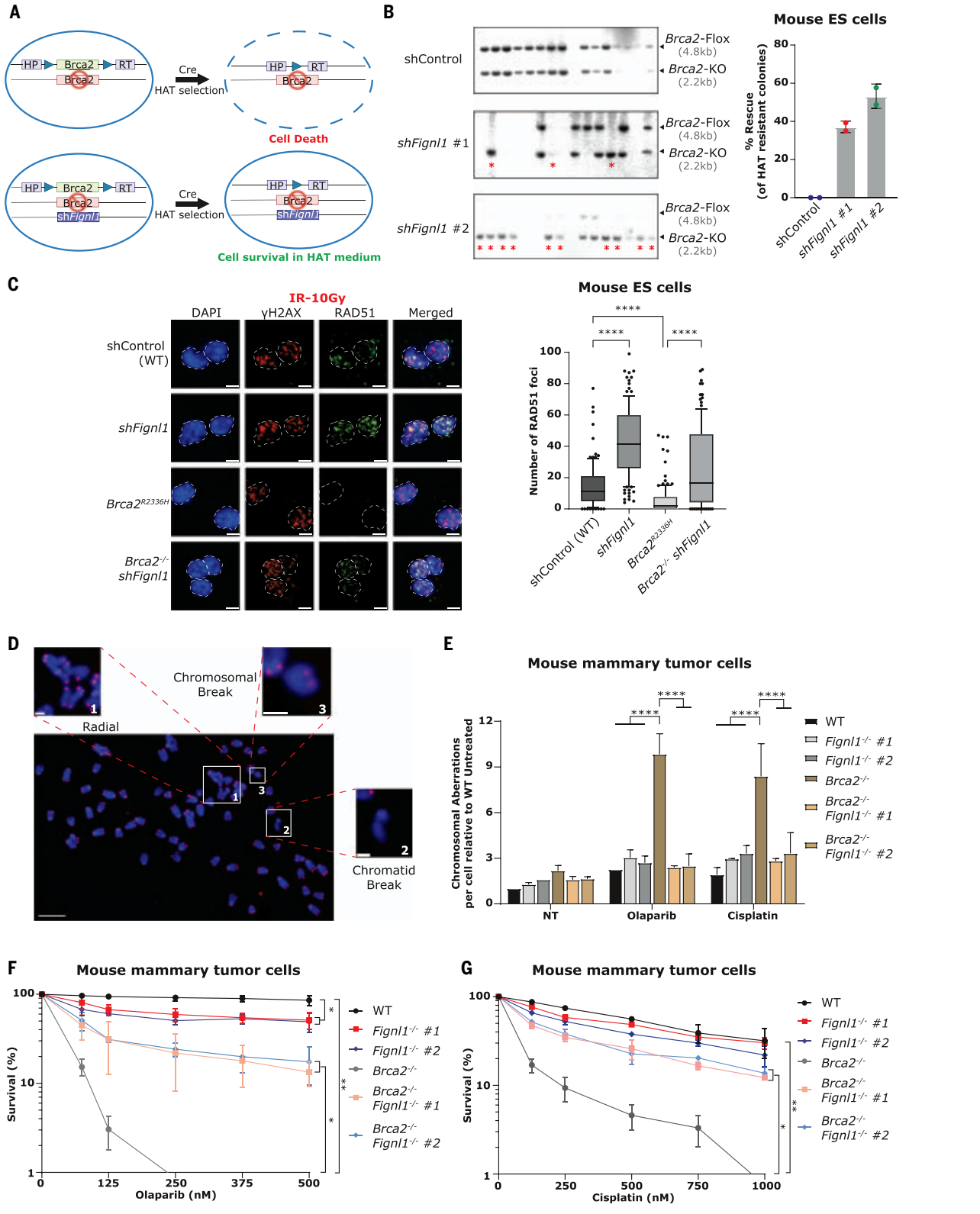


Fig. 2. FIGNL1 loss results in synthetic viability of BRCA2 null ESCs and confers resistance to chemotherapeutic drugs. (**A**) Schematic representation of strategy behind deletion of BRCA2 in mouse embryonic stem cells. (**B**) (Left) Representative southern blot images showing the 4.8 kb conditional knock-out allele (upper band) and 2.2 kb deleted allele (lower band) of *Brca2*. Stars in the shFIGNL1 clones represents the rescued clones in the absence of BRCA2. (Right) Percentage quantification of rescued clones. (**C**) (Left) Representative RAD51-IF images in mouse embryonic stem cells (scale bar: 20 μm). *Brca2*^{R2336H} a *Brca2* mutant ES cell line that expressed a hypomorphic allele (p.Arg2336His) of BRCA2 that is defective in HR. (Right) Semi-automated QIBC plots of the RAD51 foci per nucleus in WT and FIGNL1 knock down clones; scatter box plot showing the foci distribution between 10-90th percentile ($n = 150$). *P*-values from Mann-Whitney test (two-tailed) comparing between indicated groups (****: <0.0001). (**D**) Representative metaphase spread image showing different aberrations analyzed for genomic instability. Scale bar (whole metaphase): 5 μm; scale bar (individual aberrations): 1 μm.

(**E**) Quantification of chromosomal aberrations in mouse mammary tumor cells in response to Olaparib and cisplatin from three independent experiments in which 50 individual metaphases were analyzed for each sample per experiment. *P* values from ordinary two-way ANOVA (****: <0.0001). (**F**) Percentage survivability of mouse mammary tumor cells in response to Olaparib; quantification was performed by normalization to non-treated cells. *P*-values from Dunnett's multiple comparison one-way ANOVA comparing between indicated groups (*: <0.05, **: <0.01). (**G**) Percentage survivability of mouse mammary tumor cells in response to cisplatin; quantification was performed by normalization to non-treated cells. *P*-values from Dunnett's multiple comparison one-way ANOVA comparing between indicated groups (*: <0.05, **: <0.01).

FIGNL1 interacts with MMS22L-TONSL complex in the absence of RAD51

Functional RAD51 loading and HR in BRCA2/FIGNL1 double-deficient cells (Fig. 1, A, B, and F) indicated that other factors could mediate RAD51 loading and stimulate HR without BRCA2. To identify factors that could stimulate HR and be regulated by FIGNL1, we investigated protein-protein interactions (PPIs) of FIGNL1 using the PREDICTOMES classifier database (https://predictomes.org/view/ddr) (*60*). Our search revealed FIRRM, RAD51, and DMC1 as top potential interactors of FIGNL1 based on pLDDT scores (fig. S14A and table S4). FIGNL1 was predicted to interact with high confidence with both MMS22L and TONSL, which exists as a complex mediating HR though RAD51 interaction (fig. S14A and table S4) (*20*–*22*). Using AlphaFold3, we predicted interactions between monomeric FIGNL1 and MMS22L-TONSL complex. The modeled architecture showed FIGNL1 monomer wrapping around MMS22L, with NLS, FRBD, and Pore loops interacting with different MMS22L regions (Fig. 5A, fig. S15, A and B, and table S5). Although unstructured regions showed low confidence interactions, all models predicted favorable binding energy (table S5). To test the predicted PPI between FIGNL1 and MMS22L-TONSL complexes, we performed an in vitro interaction assay with purified FIGNL1-dNLS and MMS22L-TONSL (figs. S3A and S15C). The addition of MMS22L-TONSL complex to immobilized FIGNL1-dNLS revealed a direct physical interaction (Fig. 5B). To validate this in cells, we performed IP experiments using eGFP-tagged FIGNL1 in *FIGNL1*$^{-/-}$ RPE1 cells. Pulldowns of FIGNL1 contained MMS22L-TONSL complex, confirming our prediction and the in vitro data (Fig. 5C).

Since both the MMS22L-TONSL complex and FIGNL1 modulate RAD51 activity in HR, we modeled the FIGNL1-MMS22L-TONSL ternary complex with RAD51. The predicted quaternary complex suggests that with RAD51, FIGNL1 dissociates from the MMS22L-TONSL and interacts with RAD51 via the FRBD region (Fig. 5D, fig. S16, A and B, and table S6). To test this model, we performed interaction assays with purified FIGNL1-dNLS and MMS22L-TONSL complex with or without RAD51 and ssDNA-bound RAD51. RAD51-K133R (*51*, *52*) was used to ensure stronger ssDNA binding (fig. S4F). Supporting our structural prediction, RAD51-K133R reduced MMS22L-TONSL and FIGNL1 interaction without affecting RAD51 binding, with no changes observed when adding ssDNA-bound RAD51 compared to RAD51 alone (Fig. 5E and S16C).

BRCA2 has been proposed to sequester cellular RAD51 to prevent inappropriate DNA interactions (*61*). We hypothesized that RAD51 availability for the MMS22L-TONSL-FIGNL1 complex in cells could be constrained by BRCA2. Pull-down experiments of FIGNL1 showed that BRCA2 deficiency resulted in the dissociation between FIGNL1 and the MMS22L-TONSL complex compared with WT cells (Fig. 5F), potentially due to increased RAD51 availability. AlphaFold3 predictions of the FIGNL1-MMS22L-TONSL ternary complex and FIGNL1-MMS22L-TONSL-RAD51 quaternary complex indicated that FIGNL1 could interact with both RAD51 and MMS22L-TONSL through the FRBD domain, suggesting competitive interaction (Fig. 5, A and D, figs. S15, A and B, and S16, A and B, and tables S5 and S6). Our in vitro experiments showed that RAD51 interaction with FIGNL1-dNLS was lost following FRBD region deletion (Fig. 3A). Additional experiments confirmed that the absence of FRBD resulted in a loss of interaction with both RAD51 and MMS22L-TONSL, supporting that they compete for the same binding region on FIGNL1, with RAD51's primary interaction being with FIGNL1 (Fig. 5G). Subsequently, cellular pulldowns validated that FRBD loss disrupted both FIGNL1-MMS22L and FIGNL1-RAD51 interactions (Fig. 5H). BRCA2 loss, which would increase the availability of RAD51, only disrupted interaction between full-length FIGNL1 and MMS22L, whereas FRBD loss upon BRCA2 depletion disrupted both RAD51 and MMS22L interactions, further suggesting that MMS22L and RAD51 compete for the same binding site on FIGNL1 (Fig. 5H). These data suggest that the MMS22L-TONSL-FIGNL1 interaction may be critical for controlling FIGNL1 activity and limiting unwanted RAD51 removal during HR.

MMS22L-TONSL promotes RAD51 accumulation and facilitates HR in BRCA2/FIGNL1 double-deficient cells

Our data suggest that the interaction between MMS22L-TONSL and FIGNL1 regulates RAD51 dynamics upon DSB induction. We therefore assessed MMS22L localization to DSBs with or without BRCA2 and/or FIGNL1 in RPE1 and mouse mammary tumor cells. Upon IR treatments, WT cells showed increased MMS22L localization compared to untreated cells (Fig. 6A and fig. S17A). FIGNL1 loss in two clones significantly increased MMS22L foci compared with WT cells. BRCA2 loss caused a slight increase in MMS22L foci, further enhanced in BRCA2/FIGNL1 double-deficient cells (Fig. 6A and fig. S17A). The increased MMS22L upon loss of FIGNL1 was correlated with increased RAD51 localization upon IR exposure (Fig. 1A and fig. S2, A and B). These data suggest that FIGNL1 primarily dissociates RAD51 loaded by the MMS22L-TONSL complex. To validate, we conducted in vitro ssDNA protection assay with purified RAD51, the MMS22L-TONSL complex, and FIGNL1-dNLS. MMS22L-TONSL addition to RAD51-bound ssDNA in the presence of FIGNL1 did not protect against nuclease degradation (fig. S17B). Furthermore, in DNA strand exchange reactions, MMS22L-TONSL complex enhanced RAD51-mediated strand exchange (*22*) but FIGNL1-dNLS addition inhibited the reaction in a concentration-dependent manner (Fig. 6B). This indicates that the MMS22L-TONSL complex cannot counteract FIGNL1-mediated RAD51 dissociation, thus inhibiting HR.

Given the enhanced MMS22L localization in BRCA2/FIGNL1 double-deficient cells (Fig. 6A and fig. S17A), we investigated whether the MMS22L-TONSL complex facilitates RAD51 loading at DSBs. We down-regulated MMS22L, TONSL, or both in WT, *FIGNL1*$^{-/-}$ and shBRCA2 *FIGNL1*$^{-/-}$ RPE1 cells (fig. S17, C and E) and probed for RAD51 accumulation upon DSB induction. Loss of the MMS22L-TONSL complex reduced RAD51 foci in WT cells, though to a lesser extent than BRCA2 deficiency (fig. S17D). Down-regulation of the MMS22L-TONSL complex decreased RAD51 focus accumulation in *FIGNL1*$^{-/-}$ cells to nearly WT levels, supporting that FIGNL1 primarily dissociates RAD51 accumulation mediated by MMS22L-TONSL (Fig. 6C). Loss of the MMS22L-TONSL complex in BRCA2/FIGNL1 double-deficient cells almost completely abrogated RAD51 focus accumulation (Fig. 6C), suggesting that MMS22L-TONSL is essential for RAD51 accumulation in BRCA2/FIGNL1 double-deficient cells (Fig. 6C). Additionally, DR-GFP reporter assays (*56*) showed that loss of MMS22L-TONSL in BRCA2/FIGNL1 double-deficient cells significantly reduced HR efficiency (Fig. 6D and fig. S17F). These results confirm the critical role of MMS22L-TONSL in mediating HR upon the loss of BRCA2 and FIGNL1 expression. We then tested whether MMS22L-TONSL loss in BRCA2/FIGNL1 double-deficient cells affects genomic stability upon DSB-inducing treatments. Metaphase spread analysis revealed increased chromosomal aberrations upon MMS22L-TONSL complex down-regulation in sh*BRCA2 FIGNL1*$^{-/-}$ cells treated with PARPi and cisplatin (Fig. 6E). Finally, in line with the genomic instability data, clonogenic survival assays with PARPi treatment in *Brca2*$^{-/-}$ *Fignl1*$^{-/-}$ mouse mammary tumor cells showed that MMS22L loss resulted in significant cellular sensitivity compared with *Brca2*$^{-/-}$ *Fignl1*$^{-/-}$ cells

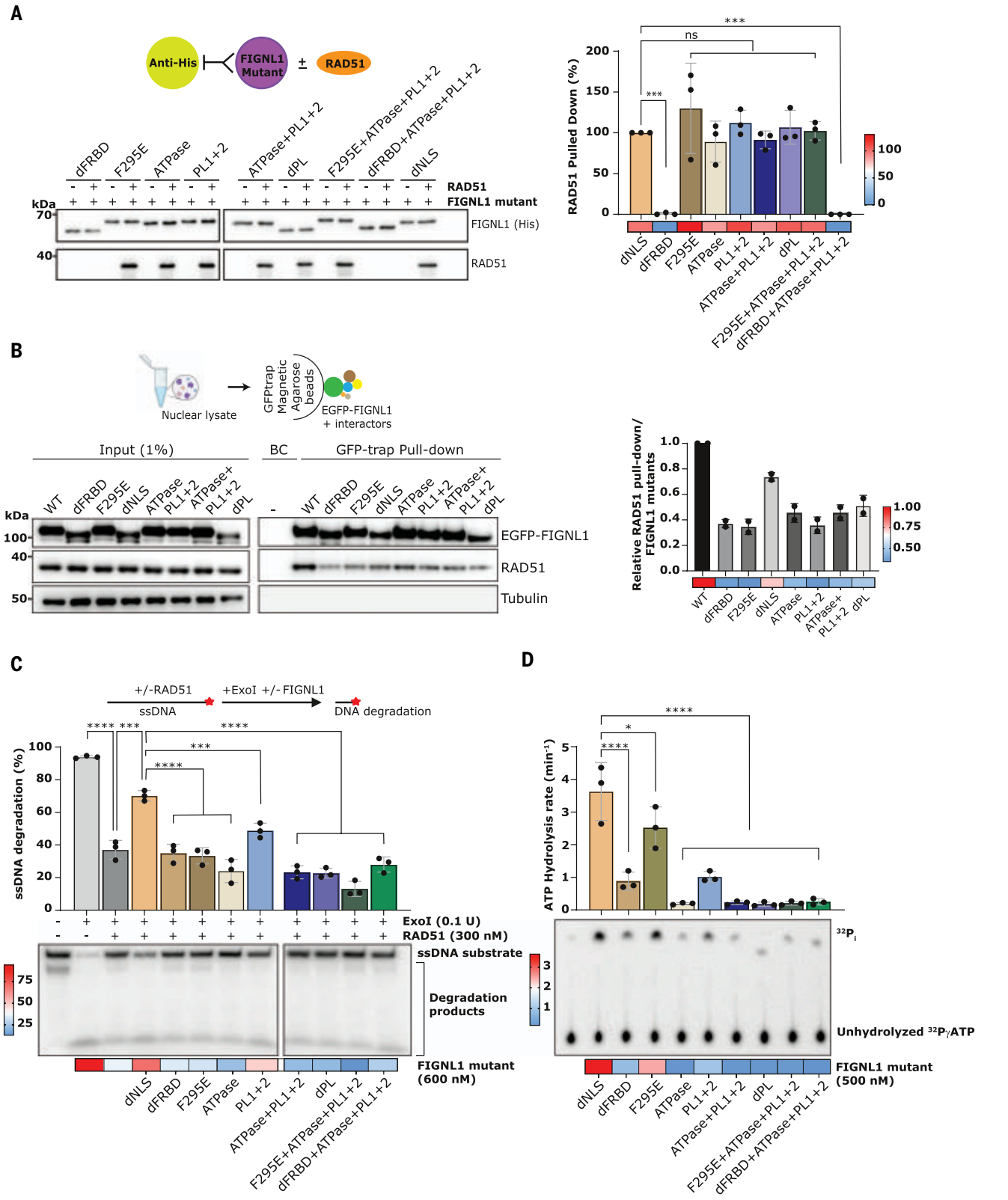


Fig. 3. FRBD and ATPase domain of FIGNL1 are important for its interaction with RAD51. (**A**) (Top) Schematics of the in vitro pull-down assay. (Left) Western blot showing the RAD51 pulled down by the different FIGNL1 mutants. (Right) Quantification of the percentage of RAD51 pulled down by the FIGNL1 mutants normalized to the dNLS mutant; error bars are indicative of mean ± SEM from three independent experiments ($n = 3$). *P*-values from Dunnett's multiple comparison one-way ANOVA comparing between indicated groups (***: <0.001). Heatmap also showing the percentage pull-down of RAD51 by the FIGNL1 mutants. (**B**) (Top) Schematics of the GFP-trap pull-down for the FIGNL1 mutants and RAD51 from the nuclear extracts. (Left) Western blot results indicative of the interaction between the FIGNL1 mutants and RAD51. (Right) Quantification of the relative RAD51 pull-down efficiency normalized for respective FIGNL1 mutants from 2 independent experiments ($n = 2$). Heatmap also showing the relative pull-down efficiency of RAD51 by the FIGNL1 mutants. (**C**) (Top) Schematics of the ssDNA protection assay showing the protection efficiency of 300 nM RAD51 against degradation by 0.1 U

ExoI nuclease in the presence or absence of 600 nM FIGNL1 mutants. The red asterisk indicates the position of the radioactive label. (Bottom) Quantification of the absolute ssDNA degradation; error bars are indicative of mean ± SEM from three independent experiments ($n = 3$). *P*-values from Dunnett's multiple comparison one-way ANOVA comparing between indicated groups (***: <0.001; ****: <0.0001). Heatmap also showing the percentage of ssDNA degradation in the presence of FIGNL1 mutants. (**D**) ATP hydrolysis rate measured in the presence of 500 nM FIGNL1 mutants. (Top) Quantification of the ATP hydrolysis rate per minute; error bars are indicative of mean ± SEM from three independent experiments ($n = 3$). *P*-values from Dunnett's multiple comparison one-way ANOVA comparing between indicated groups (*: <0.05; ****: <0.0001). Heatmap also showing the ATP hydrolysis rate in the presence of FIGNL1 mutants.

alone (Fig. 6F and fig. S17G). These data underscore the essential role of MMS22L-TONSL mediated stimulation of HR to maintain genome stability in BRCA2- and FIGNL1-deficient cells.

Discussion

BRCA2 deficiency impairs HR primarily by preventing RAD51 nucleation and filament stabilization (*5*–*11*). We show that loss of the AAA+ ATPase FIGNL1 restores RAD51 loading and HR activity, thereby conferring synthetic viability in BRCA2-deficient mESCs and resistance to PARPi inhibitors and cisplatin. We show that HR defects in BRCA2-deficient cells result from excessive RAD51 dissociation mediated by FIGNL1 (fig. S18). These findings underscore the importance of finely tuned RAD51 dynamics—both loading and dissociation—for genome maintenance and cell viability (*9*, *40*, *43*, *57*).

FIGNL1 can directly displace RAD51 from ssDNA and thus inhibit strand invasion and HR to balance the activity of BRCA2. FIGNL1 may also dissociate RAD51 from dsDNA, a process potentially relevant to reversed replication forks, in which RAD51 loading onto dsDNA is facilitated by BRCA2's C-terminal TR2 binding site for fork protection (*19*, *62*–*65*). The inability to stabilize RAD51 through TR2 interactions at reversed forks in BRCA2-deficient cells is likely attributable to unopposed FIGNL1 activity, which may result in fork degradation and genomic instability.

In BRCA2/FIGNL1-deficient cells, the MMS22L-TONSL complex emerges as the principal pathway mediating RAD51 loading and HR. Loss of this complex abrogates RAD51 accumulation and HR and increases genomic instability and drug sensitivity. The association of TONSL with H4K20me0, along with its interaction with ASF1 and CAF1, may mediate the recruitment of MMS22L-bound RAD51 to sites of DSBs on newly replicated DNA, thereby facilitating RAD51 loading in BRCA2/FIGNL1–deficient cells (*23*, *24*). The process of loading and stabilizing RAD51 in this context may also involve the interaction with other RAD51-associated proteins, including the RAD51 paralogs, which can facilitate nucleation and extension of RAD51 filaments (*66*–*68*).

Collectively, our findings reveal that inhibition of RAD51 dissociation by FIGNL1 restores HR in BRCA2-deficient cells, driving synthetic viability and drug resistance. These results have major implications for understanding mechanisms of chemoresistance in BRCA2-mutant tumors, which could restore RAD51 loading and HR without reversion mutations in the BRCA2 gene during therapy (*13*, *18*, *69*). Notably, reduced FIGNL1 expression correlates with acquired resistance to PARPi inhibitors in BRCA2-deficient mouse tumors (*70*). Targeting alternative RAD51-loading pathways, such as MMS22L-TONSL, may provide new therapeutic avenues to overcome resistance and highlights the necessity of dissecting RAD51 regulation in HR and replication stress responses.

Materials and methods

Cell Culture

Human RPE1-hTERT *TP53*$^{-/-}$*; shBRCA2* cells were cultured in DMEM/F12+GlutaMAX™ (GIBCO) media containing 10% FBS (Capricorn) and 1% penicillin-streptomycin (PS) (GIBCO) at 37°C and 5% CO_2 in a humidified incubator. Mouse mammary tumor cells KB2PR and KB2P cells [previously described (*71*)], were cultured in DMEM (Gibco) media supplemented with 10% FBS, 1% PS, 5 μg/ml insulin (Merck), 5 ng/ml murine epidermal growth factor (Merck) and 5 ng/ml cholera toxin and grown under low oxygen conditions (3% O_2, 5% CO_2 at 37°C). HEK293T, BOSC and U2-OS [both DIvA (*49*) and DR-GFP (*72*)] cells were cultured in DMEM supplemented with 10% FBS, 1% PS, 1x Non-Essential Amino Acids Solution (Gibco) and 1% Glutamax™ (Gibco) at 37°C and 5% CO_2 in a humidified incubator. All the cell lines were authenticated by short tandem repeat profiling and routinely tested for mycoplasma contamination.

Cloning and generation of stable cell line models

Human *FIGNL1* sgRNAs (oligonucleotides table below) were cloned into pDG458 (Addgene #100900) plasmid at BbsI sites on the Cas9 expression plasmid. Mouse *Fignl1* sgRNA start and stop sequences (oligonucleotides table below) were cloned into pSpCas9(BB)-2A-GFP (pX458) (Addgene #48138) at BbsI site. To generate FIGNL1 knock-out models in both human RPE1 and mouse mammary tumor cells, sgRNA cloned plasmids were transfected and FACS sorted for GFP positive cells. Sorted GFP positive cells were then seeded as single cell per well in 96 well plate. Clones grown from single cells were then genotyped and analyzed by Western blot to confirm homozygous knockouts.

FIGNL1 mutant constructs: full-length FIGNL1, RAD51 interaction mutants dFRBD (Δ295-344) and FxxA point mutant (F295E); FIRRM/FLIP interaction mutant: dNLS (Δ1-120); and ATPase point mutant (K447A+D500A) cloned in entry clones were used for LR reaction to generate expression clones using Gateway cloning (*35*). PL1+2 and dPL mutants were created by site-directed mutagenesis using the full-length WT constructs. Gateway clonase LR enzyme mix II (Invitrogen) was used to clone FIGNL1 mutants into destination vector pCW57.1-eGFP vector (pCW57.1-eGFPnls for dNLS mutant). To create vector backbones for FIGNL1 mutants, Tetracycline-Responsive promoter from pCW57.1-TRE-eGFP/eGFPnls vector was replaced with CMV promoter to create pCW57.1-CMV-eGFP/eGFPnls backbone using NEBuilder® HiFi DNA Assembly Master Mix. Finally, FIGNL1 mutant constructs were co-transfected into HEK293T cells with 3rd generation lentiviral packaging and envelop vectors to create viral particles and transduced into RPE1 FIGNL1 knock-out cells. Transduced RPE1 FIGNL1-KO cells were then sorted using GFP FACS and analyzed for eGFP fused FIGNL1 fusion protein using Western blot analysis.

KB2P-AsiSI cells were generated by transfecting doxycycline-inducible AsiSI-YFP plasmid pTRE3G-HA-ER-AsiSI-PGK-GFPVenus-IRES-rtTA3 (a kind gift from Andres Canela, Kyoto University, Japan) into BOSC cells along with pCL-Eco retro-viral packaging vector. Retro-viral particles were collected after 48 hours and transduced into KB2P cells. Transduced cells were then FACS sorted for YFP and clones were selected with equal number of AsiSI induced damage marked by 53BP1 foci levels.

For recombinant FIGNL1-dNLS protein purification, cDNA sequence of FIGNL1-dNLS was cloned into pFB-1xMBP-10xHis Baculovirus backbone or pGEX-2T backbone using NEBuilder® HiFi DNA Assembly Master Mix. FIGNL1 mutants for insect *Sf9* cell expression were created by transferring the mutant construct from pCW57.1-CMV backbone to pFB-1xMBP-10xHis Baculoviral construct using NEBuilder® HiFi DNA Assembly Master Mix.

The miBRCA2 construct was PCR amplified from the plasmid pCAGGS_Mini_BRCA2-1xFlag (*53*) (kind gift from Maria Jasin) without the Nuclear Localization Signal (NLS) using the primers Mini_BRCA2_F

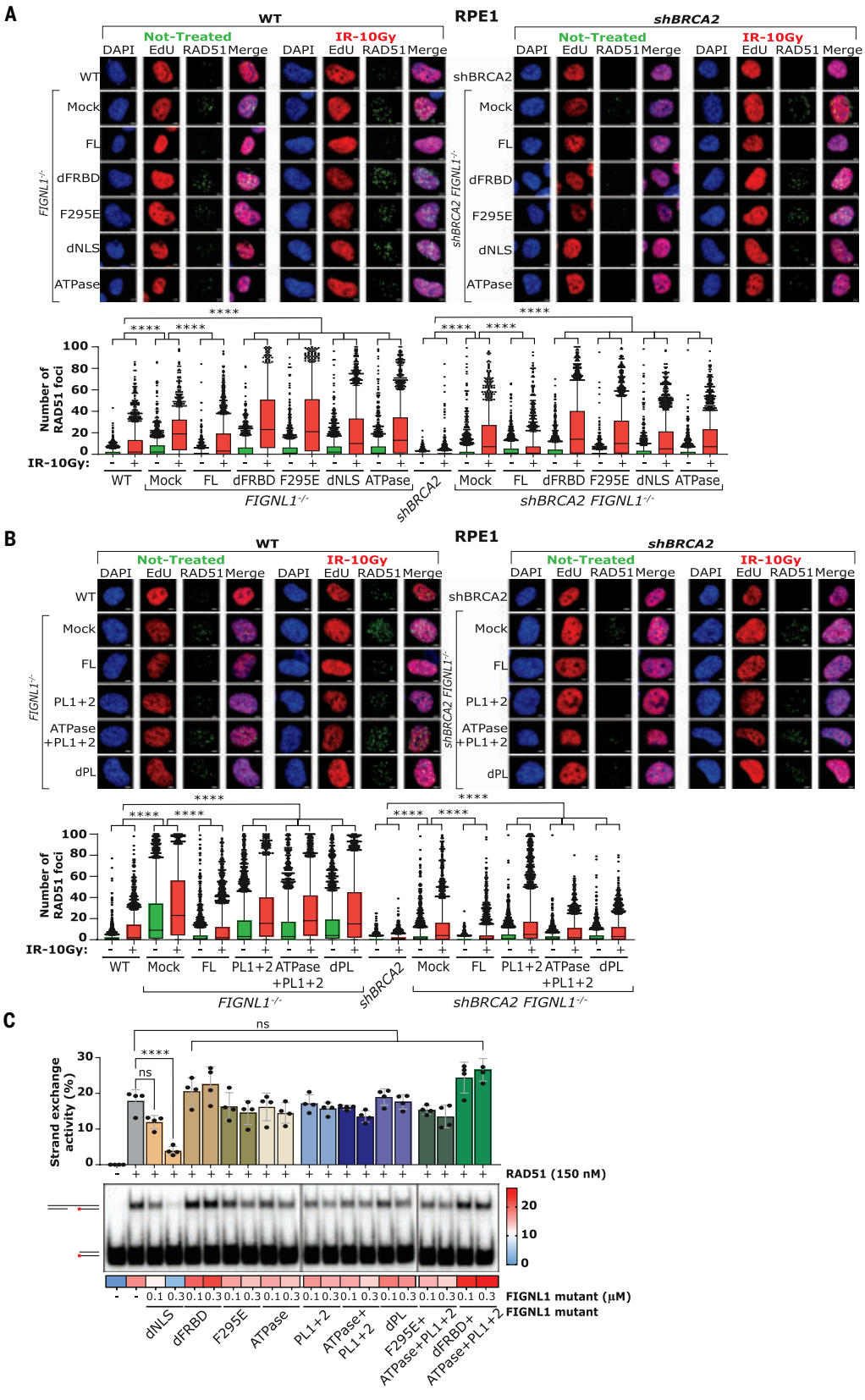


Fig. 4. FIGNL1 association with RAD51 is critical for function. (**A**) (Top) Representative images of high-content, automated RAD51-IF images showing the EdU and RAD51 foci in not-treated and ionizing radiation treated human RPE1 cells; scale bar: 5 μm. (Bottom) Automated QIBC plots of the RAD51 foci per nucleus; scatter box plot showing the foci distribution between 10-90th percentile from 1500 cells (experiment was repeated 3 times). *P*-values from Mann-Whitney test (two-tailed) comparing between respective not-treated or IR treated groups (****: <0.0001). (**B**) (Top) Representative images of high-content, automated RAD51-IF images showing the EdU and RAD51 foci in not-treated and ionizing radiation treated human RPE1 cells; scale bar: 5 μm. (Bottom) Automated QIBC plots of the RAD51 foci per nucleus; scatter box plot showing the

foci distribution between 10-90th percentile from 1500 cells (experiment was repeated 3 times). *P*-values from Mann-Whitney test (two-tailed) comparing between respective not-treated or IR treated groups (****: <0.0001). (**C**) Quantification of the absolute DNA strand exchange activity showing the band intensity of the exchanged DNA; error bars are indicative of mean ± SEM from 4 independent experiments ($n = 4$). The red asterisk indicates the position of the radioactive label. *P*-values from Dunnett's multiple comparison one-way ANOVA comparing between indicated groups (****:<0.0001). Heatmap also showing the percentage strand exchange activity of RAD51 in the presence of FIGNL1 mutants.

and Midi1_BRCA2_BRC1-2_NotI_R with introducing NheI and NotI site at the ends. The product was then cloned into the vector pFastBac-2XMBP-HLTFco-His (*73*) between the NheI and NotI sites to generate the final plasmid pFB-2xMBP-Mini_BRCA2-FLAG-His.

Transfection and infection

All the plasmid DNA transfections were performed using X-tremeGENE 9 DNA Transfection Reagent (Merck) and siRNAs were transfected using Lipofectamine RNAiMAX Transfection Reagent (Invitrogen) according to manufacturer's instructions. For viral infections, 48-hour viral soups collected from HEK293T (lenti-viral) and BOSC (retro-viral) cells were transduced into target cell lines with 8 μg/ml polybrene (Sigma-Aldrich). Viral titers were assessed by qPCR Lenti/Retrovirus Titer Kits (ABM, Canada).

ES cell culture and generation of stable lines

PL2F7 mouse embryonic stem cells (mESCs) used in the study were generated from AB2.2 mESC line by knocking out one copy of *Brca2* and flanking the other allele with two *loxP* sites along with 5′ and 3′ human *HPRT minigene* (*58*). mESCs were cultured on mitotically inactive SNL feeder cells in M15 (knockout DMEM) media (Life technologies) as described previously (*58*).

For generation of stable *Fignl1* knockdown in PL2F7 mESC, two different shRNAs against *Fignl1* and one control shRNA were cloned into PLKO lentiviral vector. Sh*Fignl1* lentiviral particles were generated by co-transfecting HEK293T cells with PLKO_sh*Fignl1* plasmid and packaging plasmid for 48 hours. PL2F7-A10 (PL2F7 clone lacking puromycin resistance gene) mESCs were transduced with PLKO_sh*Fignl1* or PLKO_Control lentiviral particles on feeder free gelatinized plate. After 48 hours, cells were trypsinized and 10^4 cells were seeded on a 10 cm plate containing SNL feeder cells. Cells were selected with puromycin (3 ug/ml) for 5 days. Visible colonies were picked and the stable knockdown of sh*Fignl1* were confirmed by Western blot.

Deletion of Brca2 cko allele and confirmation of shFignl1 Brca2$^{ko/ko}$ mESCs by Southern blot

sh*Fignl1* expressing PL2F7-A10 stable clones were used for deletion of *Brca2 cko* allele. Ten million mESCs were electroporated with 25 ug of *Pgk-Cre* plasmid. After 36 hours of electroporation, recombinant clones were selected in HAT media for 5 days followed by HT selection for 2 days. Cells were then maintained in M15 media, and mESC colonies were picked into 96 well plate for expansion and genomic DNA isolation using mESCs lysis buffer containing 10 mg/ml of proteinase K as described previously (*58*). Genomic DNA was digested with *EcoRV* at 37°C overnight and the digested DNA was electrophoresed on 1% agarose gel in 1X TBE buffer, transferred to Hybond-N+ nylon membrane. DNA probe was labelled with [α^{32}P]-dCTP, hybridized overnight, washed and imaged using Amersham Typhoon image scanner (Cytiva) as described previously (*58*). The PCR based genotyping was performed as described previously (*19*) to detect the *Brca2 cko* and *Brca2 ko* alleles, using the primers listed in the table below.

Western blot

Cells for Western blot were harvested and lysed in lysis buffer [50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 2 mM EDTA and protease inhibitor cocktail (PIC)]. The protein concentration was determined using the Pierce BCA protein Assay kit. Proteins were loaded and run on a 4 to 12% Bis-Tris pre-cast gel or a 3 to 8% Tris-Acetate pre-cast gel (Invitrogen) and were electrophoretically transferred to a PVDF membrane (Millipore). After transfer, the membranes were blocked in 5% milk for 1 hour and incubated overnight in primary antibodies. Membranes were then washed four times in 0.1% TBST and probed with horseradish peroxidase-conjugated secondary antibodies (Cytiva). Finally, the ECL Prime Western Blotting Detection Reagent kit (GE Healthcare) was used to develop the blot.

Flow-cytometry analysis of cell cycle

For flow-cytometry analysis of the cell cycle, the cells were labeled with EdU for 30 min followed by fixation in 4% formaldehyde/PBS and permeabilization in 1% saponin buffer in 1% BSA/PBS. The Click-It reaction was performed using Fluor Azide 594 according to the manufacturer's instructions (Invitrogen) to label the EdU. The cells were treated with DAPI as a nuclear stain. The samples were measured in BD LSR Fortessa and analyzed by FlowJo software v10.5.0.

High content microscopy

Immunofluorescence staining was performed as described previously (*46*). Briefly, the RPE1 and mouse mammary tumor cells were seed in 96- or 364-well plates. 48 hours (RPE1 cells) or 24 hours (mouse mammary tumor cells) after seeding the cells were treated with EdU for 20 min before irradiation with 10 Gy of x-ray. Cells were maintained at 37°C for 4 hours before the cells were pre-extracted with ice-cold CSK buffer [10 mM PIPES (pH 6.8), 100 mM NaCl, 300 mM sucrose, 1 mM EGTA, 1 mM $MgCl_2$, 1 mM DTT and PIC], fixed with 4% formaldehyde and permeabilized with 0,1% Triton X-100. The cells were blocked in 5% BSA/PBS before they underwent the Click-IT reaction with Alexa Fluor 488 (Invitrogen) according to the manufacturer's instruction. The cells were washed and then incubated with corresponding primary antibodies for 1 hour 30 min at room temperature. Next, they were incubated with Alexa Fluor-conjugated secondary antibodies (Invitrogen) for 1 hour at room temperature, after which the cells were treated with DAPI as a nuclear stain. The images were captured with the Opera Phenix High-Content spinning disc confocal microscope and analyzed using CellProfiler.

For AsiSI induced DSB-IF in KB2P-AsiSI cells, HA-tagged AsiSI was expressed by adding 2 μg/ml of doxycycline for 24 hours with last 5 hours of 400 nM 4OHT for nuclear translocation. While DivA-AsiSI cells with constitutive AsiSI expression were treated with 400 nM 4OHT after 24 hours of seeding. Cells were then processed similarly as mentioned above.

mESCs were seeded on a 24-well plate and after 24 hours, cells were irradiated at 10Gy. Cells were maintained at 37°C for 4 hours in fresh M15 media before it was fixed and stained for RAD51. Imaging was performed using Confocal spinning Disk microscopy (Opera Phenix) at 40x magnification and maximum intensity projection were used to count RAD51 foci. DAPI was used as a nuclear stain.

Automated image analysis

Z-stack images (3 stacks of 1 μm height) of individual channels from high throughput microscopy images were merged to create maximum intensity projection and subjected to CellProfiler analysis (*74*). Briefly, after the illumination correction, nuclei were identified using the DAPI channel and used as mask to identify and measure the intensities of other corresponding channels. Foci detection was performed from masked nuclear regions in the corresponding channels to automatically

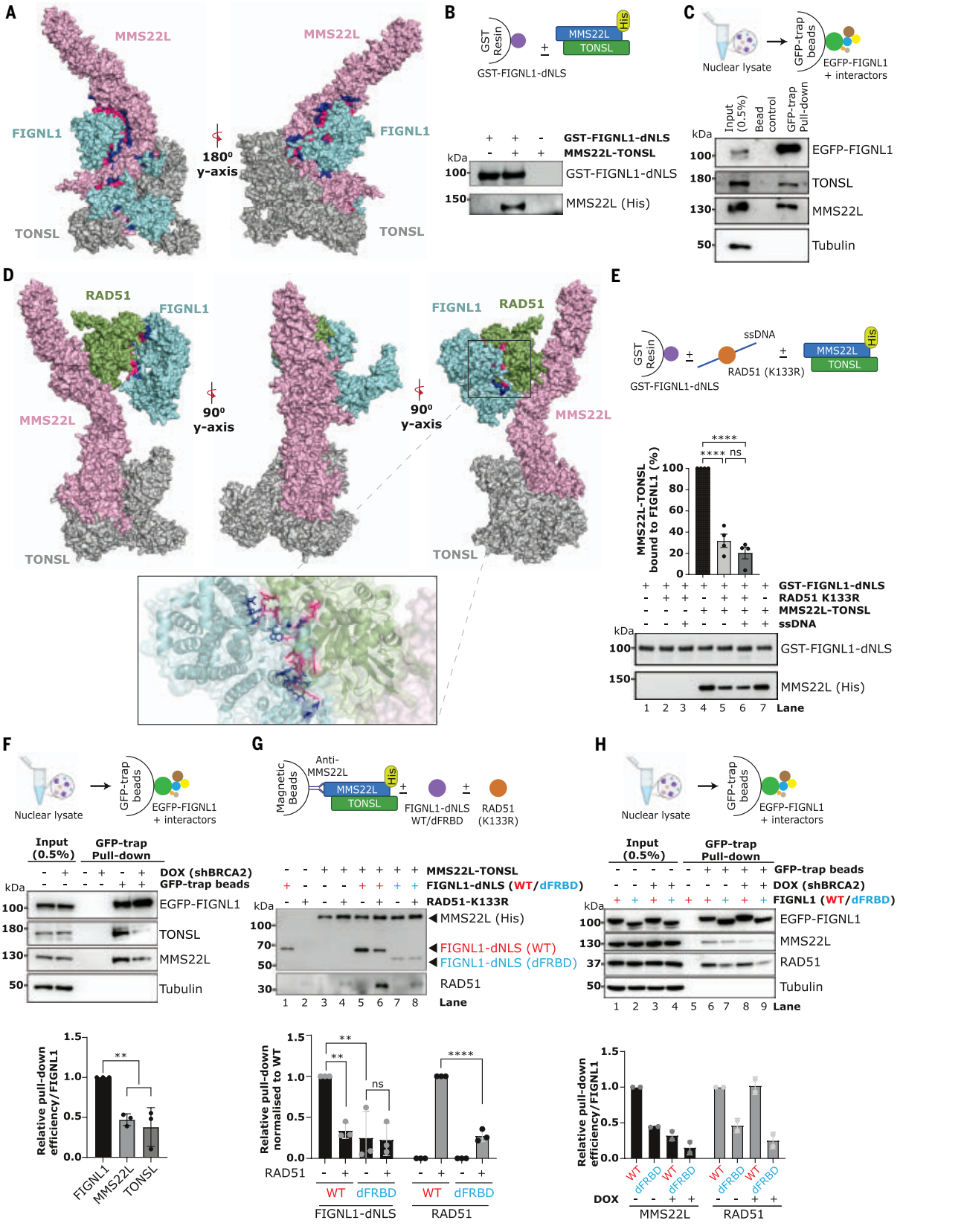


Fig. 5. RAD51 availability determines FIGNL1 interaction with MMS22L-TONSL complex via FRBD domain. (**A**) Surface representation of FIGNL1-MMS22L-TONSL ternary complex model shown in two different orientations [FIGNL1 (cyan), MMS22L (light pink), and TONSL (gray)]. The binding interface of FIGNL1 and MMS22L are depicted by pink and dark blue colors, respectively. (**B**) (Top) Schematics of the in vitro interaction assay between recombinant FIGNL1-dNLS (immobilized) and MMS22L-TONSL. (Bottom) Western blot results show that FIGNL1-dNLS interacts with MMS22L-TONSL in the in vitro interaction assay. (**C**) (Top) Schematics of the GFP-trap pull-down of the eGFP-FIGNL1 and MMS22L-TONSL from nuclear extracts. eGFP-FIGNL1 stably overexpressing cells in FIGNL1 knock-out RPE1 cells were used for nuclear fractionation and GFP-trap pulled

down. (Bottom) Western blot results indicative of the interaction between FIGNL1 and MMS22L-TONSL. (**D**) Surface representation of FIGNL1-RAD51-MMS22L-TONSL quaternary complex shown in three different orientations [FIGNL1 (cyan), RAD51 (green), MMS22L (light pink), and TONSL (gray)]. The inset shows a zoom-in view of the FIGNL1 and RAD51 interacting surface in dark blue (FIGNL1) and pink (RAD51). (**E**) (Top) Schematics of the in vitro interaction assay between recombinant FIGNL1-dNLS (immobilized), MMS22L-TONSL and RAD51-K133R in the presence or absence of ssDNA. (Bottom) Percentage quantifications showing the band intensity of the bound MMS22L-TONSL to FIGNL1-dNLS; error bars are indicative of mean ± SEM from four independent experiments ($n = 4$). *P*-values from ordinary one-way ANOVA comparing between indicated groups (****: <0.0001). Western blot results indicate the interactions between the proteins involved. (**F**) (Top) Schematics of the GFP-trap pulldown of the eGFP-FIGNL1 and its interactors from nuclear extracts. eGFP-FIGNL1 stably over-expressing cells in FIGNL1 knock-out RPE1 cells were used for nuclear fractionation and GFP-trap pulled down. (Middle) Western blot results indicative of the interaction between FIGNL1 and MMS22L-TONSL. (Bottom) Quantification showing the relative fold change of the MMS22L or TONSL binding to FIGNL1 in the presence or absence of BRCA2; error bars are indicative of mean ± SD from three independent experiments ($n = 3$). *P*-values from ordinary one-way ANOVA comparing between indicated groups (**: <0.01). (**G**) (Top) Schematics of the in vitro interaction assay between MMS22L-TONSL complex (immobilized), FIGNL1-dNLS (WT or dFRBD) and RAD51. (Middle) Western blot results indicative of the interaction between the proteins involved. (Bottom) Quantification showing the relative fold change of the FIGNL1-dNLS (WT or dFRBD) interacting with MMS22L in the presence or absence of RAD51; error bars are indicative of mean ± SEM from three independent experiments ($n = 3$). *P*-values from ordinary one-way ANOVA comparing between indicated groups (**: <0.01; ****: <0.0001). (**H**) (Top) Schematics of the GFP-trap pull-down of the eGFP-FIGNL1 and its interactors from nuclear extracts. eGFP-FIGNL1 (WT or dFRBD) stably over-expressing cells in FIGNL1 knock-out RPE1 cells were used for nuclear fractionation and GFP-trap pulled down. (Middle) Western blot results indicative of the interaction between FIGNL1, MMS22L, and RAD51. (Bottom) Quantification showing the relative fold change of the MMS22L or RAD51 interacting with FIGNL1 (WT or dFRBD), normalized to FIGNL1-WT in the presence or absence of BRCA2; error bars are indicative of mean ± SD from 2 independent experiments ($n = 2$).

calculate the intensity threshold to robustly and accurately classify the foci's local intensity maxima as foreground and separate it from the background. For unbiased detection of foci, a single threshold was set for all the groups in each experiment. For colocalization analysis between RAD51 and γH2AX, identified RAD51 foci were used as a parental object to mask on the γH2AX children object and classify the masked foci as a new group to measure the percentage colocalization. Finally, the data output was exported to TIBCO Spotfire software/Graphpad Prism for visualization.

Structural modelling of FIGNL1 molecular complexes

The full-length amino acid sequences for FIGNL1 (Q6PIW4), RAD51 (Q06609), MMS22L (Q6ZRQ5), and TONSL (Q96HA7) were obtained from UniProt. Structural models of FIGNL1 hexamer and its complexes with RAD51 in the presence of Mg^{2+}/ATP, and FIGNL1 monomer complexes with MMS22L-TONSL (ternary), and RAD51-MMSS22L/TONSL (quaternary) were generated by AlphaFold3 (*75*). Further, the FIGNL1 FRBD-RAD51 dimer structure model was generated using AlphaFold2-Multimer, ColabFold version 1.5.5. For AlphaFold2-Multimer, a single query sequence consisting of the RAD51 dimer and the FIGNL1 FRBD (295-344) amino acid sequence was provided as input using the Amber refinement model with default settings (*76*). The interacting domains and binding energies of hexameric FIGNL1 and its complex with RAD51, and monomeric FIGNL1 ternary (FIGNL1-MMS22L/TONSL), and quaternary (FIGNL1-RAD51-MMS22L/TONSL) complexes were evaluated with the help of PRODIGY (*77*). The PAE matrices and pLDDT colored structures for each predicted model were prepared using PAE viewer (*78*). All the molecular interactions were analyzed and visualized with PyMol 2.2.3 visualization software (https://legacy.ccp4.ac.uk/).

Protein purification

Human MMS22L-TONSL (*22*), FIGNL1-dNLS, miBRCA2 and RPA (*73*) were expressed in *Spodoptera frugiperda* 9 (*Sf*9) insect cells in SFX Insect serum-free medium (Hyclone) using the Bac-to-Bac expression system (Invitrogen), according to manufacturer's recommendations. MMS22L and TONSL proteins were co-expressed to prepare MMS22L-TONSL protein complex using pFB-MBP-MMS22L-His and pFB-GST-TONSL plasmids. The protein was extracted from *Sf*9 cell pellets with 325 mM NaCl as described previously (*22*). The protein was then bound to amylose resin (New England Biolabs) via the MBP tag on the MMS22L subunit. The amylose resin with bound protein was washed with Amylose Wash Buffer (50 mM Tris–HCl [pH 7.5], 2 mM β-mercaptoethanol, 200 mM NaCl, 10% glycerol, and 1 mM phenylmethanesulfonyl fluoride [PMSF]) and eluted using Amylose Wash Buffer supplemented with 10 mM maltose. Both MBP and GST tags were then cleaved by PreScission protease. The GST-tag was not used during the purification. The final protein complex was obtained by affinity purification on NiNTA agarose (Qiagen) via a 10x-His-tag on the C-terminus of MMS22L. The NiNTA resin bound protein was washed with NiNTA wash buffer (50 mM Tris–HCl [pH 7.5], 2 mM β-mercaptoethanol, 150 mM NaCl, 10% glycerol, 1 mM PMSF and 40 mM imidazole) and eluted using NiNTA wash buffer supplemented with 400 mM imidazole. The complex was then dialyzed in dialysis buffer (50 mM Tris–HCl [pH 7.5], 2 mM β-mercaptoethanol, 100 mM NaCl, 10% glycerol, 1 mM PMSF). FIGNL1-dNLS and its mutants were purified from pFB-MBP-FIGNL1-dNLS-His plasmids using identical procedures as MMS22L-TONSL, except that the Amylose wash buffer contained 1 M NaCl, but the PreScission protease digestion was done at 300 mM NaCl. FIGNL1-dNLS used for mass photometry experiments were supplemented with or without 1 mM ATP and 5 mM $CaCl_2$ in all the buffers during the purification. The miBRCA2 proteins were purified from pFB-2xMBP-Mini_BRCA2-FLAG-His plasmids using similar protocols as the MMS22L-TONSL with 1 M NaCl during the MBP washing steps, but the PreScission protease digestion was done at 300 mM NaCl. The NiNTA resin with bound miBRCA2 was washed with the NiNTA Wash Buffer containing 58 mM imidazole and eluted using NiNTA Wash Buffer supplemented with 300 mM imidazole. The final protein sample was dialyzed against storage buffer (50 mM Tris–HCl [pH 7.5], 5 mM β-mercaptoethanol, 150 mM NaCl, 10% glycerol, and 0.5 mM PMSF) (amino acid sequence of miBRCA2 is provided in the oligonucleotide table below).

Human RPA1, RPA2, and RPA3 were co-expressed to produce RPA heterotrimer using pFB-RPA1, pFB-RPA2 and pFB-6xHis-RPA3 constructs. Trimeric RPA was purified by exploiting the His-tag on RPA3 using NiNTA affinity chromatography, followed by HiTrap Blue HP, HiTrap Desalting and HiTrap Q HP chromatography columns (all GE Healthcare) using ÄKTA pure (GE Healthcare) (*73*). Wild type RAD51, as well as the RAD51-K133R variant, were expressed from pMALT-P-RAD51 and pMALT-P-RAD51-K133R plasmids in BL21(DE3)pLysS *E. coli* cells and purified using amylose affinity chromatography (New England Biolabs) followed by Hitrap Q HP chromatography (Cytiva) (*22*). The MBP tag was cleaved during the purification of all RAD51 variants. Human DNA2 was expressed from the pFB-His-DNA2-FLAG plasmid in *Sf*9 insect cells and purified by affinity chromatography taking advantage of the N-terminal 6x His-tag and the C-terminal FLAG-tag (*79*). For purification of human RAD54, *Sf*9 cells were infected with virus produced with the pFB_MBP-TEV-hRAD54co-FLAG vector, and the protein was purified by amylose and FLAG affinity purification as described before (*80*).

For purification of human mitochondrial single-stranded DNA binding protein (mtSSB), pMALT-P-mtSSB vector was transformed into in *E. coli* BL21 cells and mtSSB was purified with amylose affinity purification

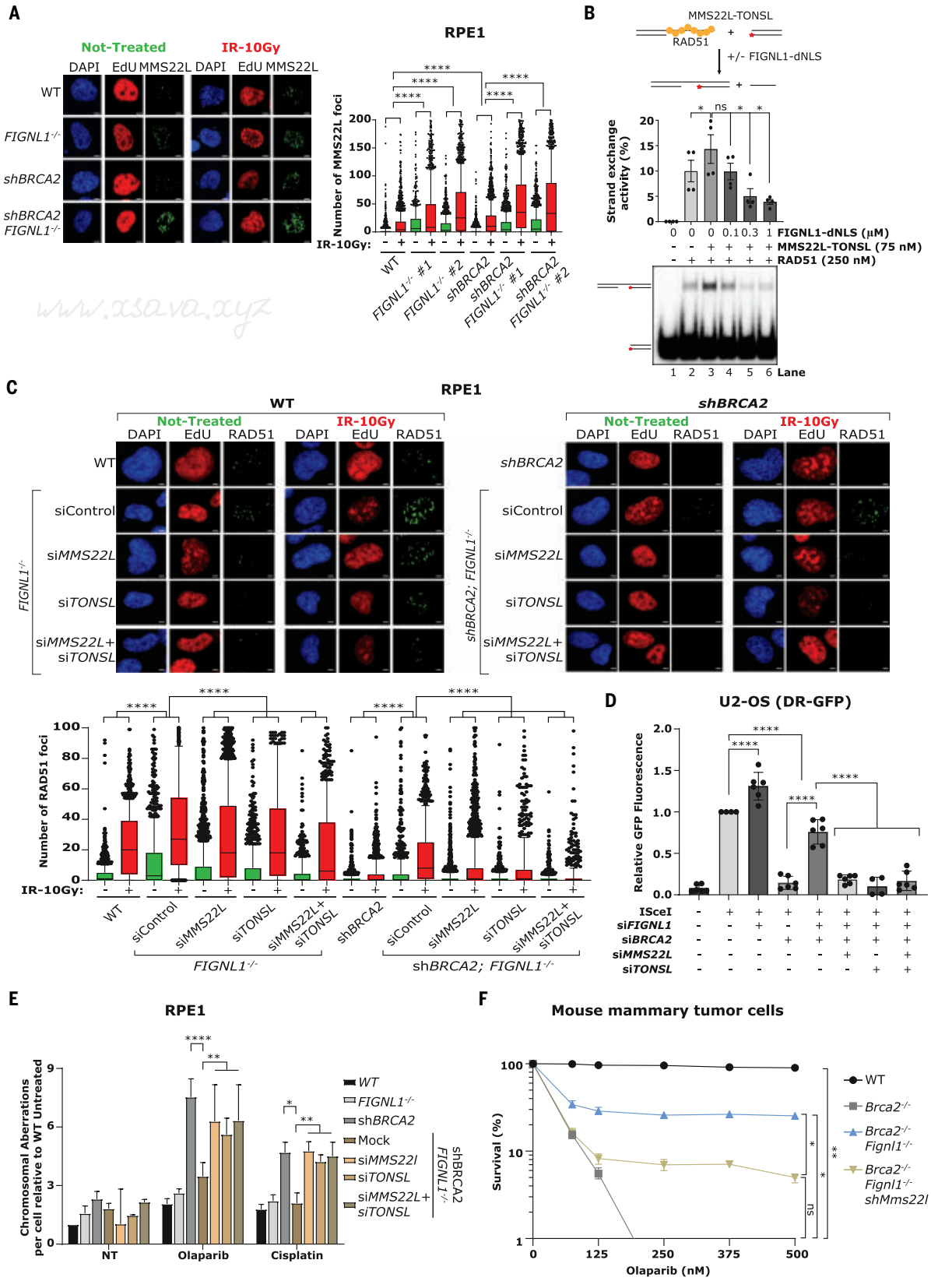


Fig. 6. MMS22L-TONSL complex loads RAD51 in the BRCA2-FIGNL1–deficient cells to recue HR and confer chemoresistance. (**A**) (Left) Representative high-content, automated IF images showing EdU and MMS22L foci in not-treated and ionizing radiation treated cells; scale bar, 5 μm. (Right) Automated QIBC plots of the MMS22L foci per nucleus; scatter box plot showing the foci distribution between 10th and 90th percentiles from 2,000 cells (experiment was repeated three times). *P*-values from Mann-Whitney test (two-tailed) comparing between respective not-treated or ionizing radiation treated groups (****: <0.0001). (**B**) (Top) Schematics of the DNA strand exchange assay. The strand exchange assays were performed by adding 250 nM purified RAD51 and 75 nM MMS22L-TONSL to the DNA substrate. The assay was then supplemented by increasing

concentrations of purified FIGNL1-dNLS. The red asterisk indicates the position of the radioactive label. (Bottom) Quantification of the absolute DNA strand exchange activity showing the band intensity of the exchanged DNA; error bars are indicative of mean ± SEM from four independent experiments ($n = 4$). *P*-values from ordinary one-way ANOVA comparing between indicated groups (*: <0.05). (**C**) (Top) Representative high-content, automated IF images showing EdU and RAD51 foci in not-treated and ionizing radiation treated human RPE1 cells; scale bar, 5 μm. (Bottom) Automated QIBC plots of the RAD51 foci per nucleus; scatter box plot showing the foci distribution between 10th and 90th percentiles from 1,500 cells (experiment was repeated three times). *P*-values from Mann-Whitney test (two-tailed) comparing between indicated groups (****: <0.0001). (**D**) DR-GFP system stably integrated U2-OS cells transfected for 72 hours with indicated siRNAs and 48 hours of ISceI to measure the relative levels of GFP fluorescence to the WT cells. Error bars are representative of the mean ± SD from three independent experiments ($n = 3$). *P*-values from unpaired *t*-test comparing between indicated groups (****: <0.0001). (**E**) Quantification of chromosomal aberrations in RPE1 FIGNL1 knockout clones with siRNAs against MMS22L and TONSL in response to Olaparib and cisplatin from three independent experiments in which 50 individual metaphases were analyzed for each sample per experiment. *P*-values from ordinary two-way ANOVA (*: <0.05; **: <0.01; ****: <0.0001). (**F**) Percentage survivability of mouse mammary tumor cells in response to Olaparib; quantification was performed by normalization to non-treated cells. *P*-values from Dunnett's multiple comparison one-way ANOVA comparing between indicated groups (*: <0.05, **: <0.01).

followed by HiTrap Blue column (GE Healthcare) as previously described (*79*).

DNA substrate preparation

For DNA strand exchange assays (*22*), the 3′-tailed DNA was prepared by annealing oligonucleotides RJ-167-mer and RJ-PHIX-42-1 in 1:1 ratio with the annealing buffer (100 mM Tris-HCl [pH 8.0], 500 mM NaCl and 100 mM $MgCl_2$). The dsDNA was generated by radio-labeling RJ-Oligo1 with ^{32}P at the 3′-end and annealing it to RJ-Oligo2 in 1:2 ratio. DNA protection assays were performed using PC1253 as ssDNA oligonucleotide. DNA binding assays were performed using PC1253 ssDNA oligonucleotide annealed with PC1253c complementary oligonucleotide to form the dsDNA. All oligonucleotides were ^{32}P-labeled at the 3′ terminus with [α-^{32}P] dCTP (Hartmann Analytic) and terminal transferase (New England Biolabs) according to the manufacturer's instructions. Sequences of all the oligonucleotide substrates used in the study are listed in the table below.

DNA binding assay

Binding reactions (15 μl volume) were carried out in binding buffer (25 mM Tris- acetate [pH 7.5], 3 mM EDTA, 1 mM dithiothreitol [DTT], and 100 μg/ml bovine serum albumin [BSA]), and indicated DNA substrate (1 nM, in molecules). 3 mM EDTA was replaced with 1 mM ATP (Sigma, A7699) and 2 mM magnesium acetate, where indicated. Proteins were added and incubated for 15 min on ice. Loading dye (5 μl; 50% glycerol [w/vol] and bromophenol blue) was added to the reactions and the products were separated on 6% polyacrylamide gels (ratio acrylamide:bisacrylamide 19:1, Bio-Rad) in TAE (40 mM Tris, 20 mM acetic acid and 1 mM EDTA) buffer at 4°C. The gels were dried on 17 CHR paper (Whatman), exposed to a storage phosphor screen (GE Healthcare) and scanned by a Typhoon Phosphor Imager (FLA9500, GE Healthcare). Signals were quantified using ImageJ and plotted with GraphPad Prism. The binding assays for the FIGNL1-dNLS mutants were carried out similarly, with the exception that the ssDNA substrate was used at a concentration of 0.5 nM in molecules.

DNA strand exchange assay

2 nM (in molecules) cold 3′ tailed DNA substrate was incubated with the indicated concentration of RAD51 for 5 min in a buffer (25 mM Tris Acetate [pH 7.5] 1 mM DTT, 1 mM $MgCl_2$, 2 mM ATP/AMP-PNP, 2 mM $CaCl_2$, and 0.1 mg/ml Recombinant Albumin [NEB]) in a total reaction volume of 15 μl. FIGNL1-dNLS was titrated on top and incubated at 37°C for another 5 min. This was followed by adding 2 nM ^{32}P-radiolabeled dsDNA (in molecules) and incubation for 30 min at 37°C. The reaction was stopped with 5 μl of stop solution (150 mM EDTA, 0.2% sodium dodecyl sulfate, 30% glycerol and bromophenol blue) and 1 μl Proteinase K (Roche, 18 mg/ml) supplemented with 10-fold excess of cold RJ oligo 1 for another 10 min at 37°C. Reaction products were separated by running 6% PAGE gel in 1X TAE buffer for 55 min at 8 V/cm. For reactions with RPA, 50 nM RPA was preincubated with the cold substrate for 5 min at 37°C before adding RAD51. For reactions with miBRCA2, 30 nM of miBRCA2 was incubated with RAD51 for 5 min at 37°C. Where indicated, MMS22L-TONSL was incubated after adding RAD51 for 5 min at 37°C before adding FIGNL1-dNLS. Where proteins were omitted, protein dilution buffer was (25 mM Tris- HCl [pH 7.5] and 100 mM NaCl) used instead. The gels were dried on 17 CHR paper (Whatman), exposed to a storage phosphor screen (GE Healthcare) and scanned by a Typhoon Phosphor Imager (FLA9500, GE Healthcare). Signals were quantified using ImageJ and plotted with GraphPad Prism.

DNA protection assays

DNA2 and Exonuclease I (*E. coli,* New England Biolabs)-catalyzed nuclease assays were performed in a 15 μl volume reaction buffer (25 mM Tris-acetate [pH 7.5], 2 mM magnesium acetate, 1 mM ATP, 1 mM DTT, 0.1 mg/ml BSA, 1 mM phosphoenolpyruvate [PEP], and 80 U/ml pyruvate kinase [Sigma]), and 7.5 nM DNA substrate (in molecules). RAD51-WT (as indicated in the figure) was preincubated with the substrates for 15 min at 37°C in the reaction buffer, which was then supplemented with the indicated amounts of DNA2 or Exonuclease I, FIGNL1-dNLS and MMS22L-TONSL and the reaction was continued for 30 min at 37°C. Reactions were stopped by adding 0.5 μl of 0.5 M EDTA and 1 μl Proteinase K (Roche, 18 mg/ml), and incubated at 50°C for 30 min. An equal amount of formamide dye (95% [v/v] formamide, 20 mM EDTA, bromophenol blue) was added, samples were heated at 95°C for 4 min and separated on 15% denaturing polyacrylamide gels (ratio acrylamide:bis acrylamide 19:1 [Biorad]). After fixing in a solution (40% methanol, 10% acetic acid and 5% glycerol) for 30 min, the gels were dried on 3MM paper (Whatman), exposed to storage phosphor screens (GE Healthcare) and scanned with a Typhoon Phosphor Imager (FLA9500).

DR-GFP assay

U2-OS cells with stable integration of DR-GFP were transfected with respective siRNAs for 24 hours followed by ISceI expressing pCBASceI transfection for further 48 hours. Cells were then harvested, washed with PBS+10% FBS and filtered through 40 μm cell strainer to prepare single cells suspension. Samples were then measured for GFP in BD LSR Fortessa and analyzed by FlowJo software v10.5.0. Remaining samples were then lysed and analyzed by Western blotting for respective protein knockdowns.

D-loop assay

26.6 nM of labelled 89-mer ssDNA (D-loop-89) was incubated with 800 nM of RAD51 for 15 minutes at 37°C in reaction buffer (25 mM Tris Acetate [pH 7.5], 1 mM DTT, 3 mM $MgCl_2$, 2 mM ATP, 2 mM $CaCl_2$, 2 mM PEP, 80 U/ml pyruvate kinase, and 0.1 mg/ml Recombinant Albumin [NEB]) in a total volume of 15 μl. FIGNL1-dNLS was titrated on top and incubated for another 5 min at 37°C. This was followed by addition of 25 nM of pUC19 plasmid (*81*) and 150 nM of RAD54 and incubated for another 20 min at 37°C. The reaction was stopped with 5 μl of stop solution (150 mM EDTA, 2% sodium dodecyl sulfate, 30%

glycerol and bromophenol blue) and 1 µl Proteinase K (Roche) for another 15 min at 37°C. Reaction products were separated on 1% agarose gel in 1x TAE buffer for 75 min. The gel was dried on DE81 paper, exposed to storage phosphor screens (GE Healthcare) and scanned with a Typhoon Phosphor Imager (FLA9500). Signals were quantified using ImageJ and plotted with GraphPad Prism.

Genome stability assay

The genome stability assay was carried out according to the standard protocol described previously (*82*). Cells were treated with 500 nM Olaparib for 24 hours or 750 nM cisplatin for 6 hours. Cells that were treated with Olaparib were exposed to colcemid containing media 12 hours after the start of the treatment and continued the Olaparib and colcemid treatment for the next 12 hours. For the cells that were treated with cisplatin, the drug treated media was washed off and the cells were exposed to colcemid containing media for the next 12 hours. Metaphase spreads were prepared by conventional methods. Metaphase slides were washed with 2x SSC, PBS and $MgCl_2$ in PBS, denatured at 80°C after which telomere probes were hybridized onto the metaphase spreads. After hybridization, the slides were washed with 50% formamide in 2x SSC, 0,1X SSC and 4x SSC containing 0,1% Tween-20. The metaphase spreads were stained with DAPI and mounted with MOWIOL. A minimum of 50 metaphase images were captured using Metafer5 and analyzed with Adobe Photoshop for chromosomal aberrations.

Survival assay

For the mESCs, 10,000 cells were seeded per well in a gelatinized 96-well plate. After 24 hours, cells were treated with varying concentrations of Olaparib or cisplatin for 72 hours. After 72 hours, cells were rinsed with PBS and incubated with 1 mg/ml of XTT (2,3-bis(2-methoxy-4-nitro-5-sulfophenyl)-5-[(phenylamino)carbonyl]-2H-tetrazolium hydroxide), 2 mM phenazine methosulfate (PMS) at 37°C for 60 min. Plates were read in iMark Microplate reader (Bio-Rad) at 455 nm.

1×10^3 mouse mammary tumor cells were seeded in a 6-well plate and 24 hours later cells were treated with indicated concentrations or Olaparib or cisplatin. Cells treated with Olaparib were refreshed with fresh Olaparib every 48 hours for 7 days, while cells treated with cisplatin were refreshed after 12 hours and incubated in fresh media without cisplatin for 7 days. After 7 days, survivability was quantified with CellTiter-Blue® Cell Viability Assay (Promega) as per manufacturer's instructions and plates were fixed and stained with Brilliant Blue G (Sigma Aldrich).

Interaction assays

To examine the interaction between the purified FIGNL1 mutants with RAD51, 1 µg of anti-HIS mouse antibody (MBL, D291-39, mouse) was immobilized on 10 µl protein G magnetic beads (Dynabeads, Invitrogen). The antibodies were incubated in 50 µl PBS-T (PBS containing 0.1% Tween-20) for 1 hour at 4°C with continuous rotation. After incubation, the beads were washed three times with 150 µl of wash buffer (50 mM Tris-HCl [pH 7.5], 3 mM EDTA, 1 mM DTT, 100 mM NaCl, 0.05% Triton X-100 [Sigma]). The beads were then mixed with 1 µg of each FIGNL1 mutant proteins in 60 µl of binding buffer (25 mM Tris-HCl [pH 7.5], 3 mM EDTA, 1 mM DTT, 0.2 µg/µl BSA, 100 mM NaCl) and incubated at 4°C with continuous rotation for 1 hour. Beads were again washed 3 times with 150 µl of wash buffer and incubated with 1 µg of RAD51 followed by incubation for another 1 hour at 4°C in continuous rotation. Beads were again washed 3 times with 200 µl of wash buffer and the proteins were eluted by boiling the beads in SDS buffer (50 mM Tris-HCl [pH 6.8], 1.6% sodium dodecyl sulfate, 1 mM DTT, 10% glycerol, 0.01% bromophenol blue) for 3 min at 95°C. The eluate was separated on a 7.5% SDS-PAGE gel, and the proteins were detected by Western blot using Anti-HIS (rabbit) and anti-RAD51 (rabbit) for FIGNL1 mutants and RAD51 respectively.

To study the interaction between FIGNL1-dNLS and MMS22L-TONSL in the presence of RAD51-K133R mutant, soluble extracts of GST-FIGNL1-dNLS prepared from pGEX-2T-FIGNL1-dNLS plasmid in Rosetta (DE3) pLysS *E. coli* cells were used. Each culture was supplemented with 0.2% glucose, induced with 0.5 mM IPTG and grown overnight at 18°C. The cells were then pelleted at 2500 g for 15 min at 4°C, washed once with STE buffer (10 mM Tris-HCl [pH 8], 500 mM NaCl, 1 mM EDTA), snap-frozen and kept at -80°C until use. The soluble extract containing GST-FIGNL1-dNLS was prepared by resuspending the bacterial pellets in GST-buffer (50 mM Tris HCl [pH 7.5], 5 mM $MgCl_2$, 2 mM ATP, 1 mM PMSF, 2 mM β-mercaptoethanol, 150 mM NaCl, 10% glycerol, 1 mM EDTA) followed by sonication and centrifugation for 30 min at 48,000 g. The supernatant was immobilized on GST resin (GE Healthcare) upon incubation for 1 hour at 4°C with continuous rotation. The resin was washed 3 times with GST-buffer post incubation. Next, either 50 µl binding buffer (50 mM Tris–HCl, [pH 7.5], 2 mM β-mercaptoethanol, 3 mM EDTA, 100 mM NaCl, 0.2 µg/ml bovine serum albumin [BSA], 1 mM PMSF) alone or additionally supplemented with 0.5 µg of RAD51-K133R or RAD51-K133R (pre-bound with ssDNA [PC1253]) was added to the immobilized GST-FIGNL1-dNLS for 1 hour at 4°C with continuous rotation. Where indicated, ssDNA was pre-bound for 1 hour to RAD51-K133R at 37°C. Next, the resin was washed once with GST buffer and MMS22L-TONSL was added to the appropriate samples and incubated again for 1 hour, followed by three washes and elution using GST buffer supplemented with 10 mM reduced glutathione (Sigma). The eluates were separated on a 10% SDS-PAGE gel and proteins were detected by Western blotting.

For the interaction between FIGNL1-dNLS WT and dFRBD with MMS22L-TONSL, 0.1 µg of anti- MMS22L antibody was immobilized on 10 µl protein G magnetic beads (Dynabeads, Invitrogen). The antibodies were incubated in 50 µl PBS-T (PBS containing 0.1% Tween-20) for 1 hour at 4°C with continuous rotation. After incubation, the beads were washed 3 times with 150 µl of PBS-T. The beads were then mixed with 1 µg of MMS22L-TONSL in 60 µl of binding buffer (25 mM Tris-HCl [pH 7.5], 3 mM EDTA, 1 mM DTT, 0.2 µg/µl BSA, 100 mM NaCl, 2 mM ATP, 5 mM $MgCl_2$) and incubated at 4°C with continuous rotation for 1 hour. Beads were washed 3 times with 150 µl of wash buffer (50 mM Tris-HCl [pH 7.5], 3 mM EDTA, 1 mM DTT, 150 mM NaCl, 0.05% Triton X-100 [Sigma], 2 mM ATP, 5 mM $MgCl_2$) and incubated with 1 µg of FIGNL1-dNLS WT or equimolar amount of the dFRBD mutant, followed by incubation for another 1 hour at 4°C in continuous rotation. Beads were again washed 2 times with 150 µl of Wash Buffer. The beads were then again incubated with 1 µg of RAD51-K133R for another 1 hour at 4°C. Lastly, the beads were washed 3 times with 150 µl of wash buffer and the proteins were eluted by boiling the beads in SDS buffer (50 mM Tris-HCl [pH 6.8], 1.6% sodium dodecyl sulfate, 1 mM DTT, 10% glycerol, 0.01% bromophenol blue) for 3 min at 95°C. The eluate was separated on a AccuPAGE Tris-Glycine Precast PAGE Gel 4-20% (Bio Bench) gel, and the proteins were detected by Western blot.

Silver staining determination of stoichiometry

To determine the FIGNL1-RAD51 stoichiometry, an interaction assay between FIGNL1 and RAD51 at various molar ratios was performed as described above for the FIGNL1 mutants. The samples were run on a 7.5% SDS-PAGE gel and analyzed by silver staining using standard procedure.

Mass photometry

Borosilicate microscope glass plate (No. 1.5 H thickness, 24 × 50 mm, VWR) were washed by soaking in Milli-Q-water, followed by isopropanol and, lastly, Milli-Q-water again. The slides were dried under a stream of clean nitrogen. Silicone gaskets (CultureWell Reusable Gasket, Grace Bio-Labs) were placed on the clean coverslip to create a well for sample delivery. For calibration and determination of molecular weights, a protein size marker (NativeMark Unstained Protein Standard, Invitrogen) was measured on the same day. All experiments were carried out with a

Oligonucleotides used in the study

Sl. No.	Name	Sequence	Organism	Purpose
1	sgRNA1	TGGCATATGTACCGGACCGA	Human	CRISPR-KO
2	sgRNA2	TCGGCCACAAGAAATTGATG	Human	CRISPR-KO
3	sgRNA-Start	CTAAAATGCAGACCTCCAGC	Mouse	CRISPR-KO
4	sgRNA-Stop	TTTGGTTGTGGAAAGTAAGT	Mouse	CRISPR-KO
5	shFignl1 #1 (mESCs)	CTTGACTACTGACTGATATAT	Mouse	shRNA
6	shFignl1 #2 (mESCs)	GTTCGACCAATAGCTTATATT	Mouse	shRNA
7	hFIGNL1-GT-For	GTCCAGTGTCCCTTCCTTCA	Human	Genotyping
8	hFIGNL1-GT-Rev	ACACATCCCCAACTTTCTATGC	Human	Genotyping
9	mFignl1-GT-For	GCATCACATTTCCCAAGTCCT	Mouse	Genotyping
10	mFignl1-GT-Rev	GAACCTCAAAGGCGGTGGG	Mouse	Genotyping
11	Brca2-KO-For	GCCACCTCTGCCTGATTCTA	Mouse	Genotyping
12	Brca2-KO-Rev	AAAGAACCAGCTGGGGCTCGAG	Mouse	Genotyping
13	Brca2-flox-For	TGAAGTGGACCCTGTAACCC	Mouse	Genotyping
14	Brca2-flox-Rev	AGTTCTCTCCTTTCAGCCTTCT	Mouse	Genotyping
15	CMV_EGFP_Frag.FOR	TTTGGCCGCGAATCGATATGCGTTACATAACT	Human	Gibson assembly
16	CMV_EGFP_Frag.REV	CATGgctagCCAATTCTCCAAGCTCTGCTTATATAGACCTCCCAC	Human	Gibson assembly
17	CMV_EGFP_Vector.FOR	AGGTCTATATAAGCAGAGCTTGGAGAATTGGctagcCATGG	Human	Gibson assembly
18	CMV_EGFP_Vector.REV	TTACCGTAAGTTATGTAACGCATATCGATTCGCGGCCAAAGT	Human	Gibson assembly
19	pFB-FIGNL1_Frag.FOR	aagttctgttccaggggcccTCCATGGCTGGCAAAAAAT	Human	Gibson assembly
20	pFB-FIGNL1_Frag.REV	tgatggtgctcgagcccgggCTTTCCACAACCAAAAGTTTTGTTCCAG	Human	Gibson assembly
21	pFB-FIGNL1-Vector.FOR	AAACTTTTGGTTGTGGAAAGcccgggctcgagcac	Human	Gibson assembly
22	pFB-FIGNL1-Vector.REV	AATTTTTTGCCAGCCATGGAgggcccctggaacagaacttccagc	Human	Gibson assembly
23	FIGNL1-dNLS_pGEX-For	atctggttccgcgtggatccTCCATGGCTGGCAAAAAATTCAAAG	Human	Gibson assembly
24	FIGNL1-dNLS_pGEX-Rev	gtcagtcacgatgaattctaTTACTTTCCACAACCAAAAGTTTTGTTCC	Human	Gibson assembly
25	Vector_FIGNL1-pGEX-For	CTTTTGGTTGTGGAAAGTAAtagaattcatcgtgactgactgacga	Human	Gibson assembly
26	Vector_FIGNL1-pGEX-Rev	AATTTTTTGCCAGCCATGGAggatccacgcggaaccag	Human	Gibson assembly

TwoMP mass photometer (Refeyn Ltd). Reactions containing FIGNL1-dNLS with or without RAD51 were prepared in reaction buffer containing 30 mM HEPES-KOH (pH 7.5), 100 mM KCl, 20 mM NaCl, 3.5 mM $MgCl_2$, 5% glycerol and 0.8 mM ATPγS (BIOLOG Life Science Institute). The protein concentrations for initial binding were 1 μM to allow for protein complex formation. The samples were incubated for 10 min at 37°C, followed by 1 hour at room temperature. The samples were then diluted for measurement to 100 nM using a dilution buffer containing 25 mM Tris-HCl (pH 7.5) and 100 mM NaCl in a final volume of 20 μl. Movies were recorded for 120 s using AcquireMP (Refeyn Ltd) and processed and analyzed using DiscoverMP (Refeyn ltd).

ATPase assay

ATPase assays with wild-type and mutant FIGNL1-dNLS were performed in reaction buffer (25 mM Tris-acetate [pH 7.5], 1 mM magnesium acetate, 1 mM DTT, 0.1 mg/ml BSA, 0.1 mM ATP, 1 nM of [γ-32P] ATP [Hartmann-Analytic]) in 10 μl reaction volume. Recombinant proteins were added as indicated. The samples were mixed on ice and incubated at 37°C for 30 min. Reactions were stopped with 1.25 μl of 0.5 M EDTA and separated using thin-layer chromatography plates (Merck) with 0.3 M LiCl and 0.3 M formic acid as the mobile phase. Dried plates were exposed to storage phosphor screens (GE Healthcare) and scanned by a Typhoon Phosphor Imager (FLA9500). Signals were quantified using ImageJ software. Spontaneous ATP hydrolysis signal from no protein lanes was removed as a background during quantitation. The rate of ATP hydrolysis was obtained as a measure of the amount of hydrolyzed ATP per molecule of FIGNL1-dNLS per min.

GFPtrap pull down

$10x10^6$ eGFP-FIGNL1 over expressing RPE1 cells were harvested and subjected to nuclear fractionation. Briefly, cells were lysed with cytoplasmic extraction buffer [10 mM HEPES (pH-8.0), 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, 0.05% NP40 supplemented with protease inhibitor cocktail (PIC)] and incubated on ice for 10 min. separated by centrifugation at then separated by centrifugation at 1400 g for 5 min. Nuclear fraction at the pellet was then separated and lysed with immunoprecipitation buffer (50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 10% glycerol, 1% Triton X-100, 0.5 mM EDTA, 0.5 mM EGTA, and Protease Inhibitor cocktail) supplemented with 50 U/ml Benzonase, on ice for 30 min. Finally, the lysed nuclear fraction was collected by centrifugation at 17,000 g for 10 min. Cleared nuclear fraction was quantified by BCA method (Pierce, Invitrogen) and used for overnight GFPtrap magnetic agarose bead pull-down (Proteintech) after separating 5% lysate as input. After overnight GFPtrap pull down, beads were washed 5 times with wash buffer (10 mM Tris-HCl [pH 7.4], 150 mM NaCl, 0.05% NP40 Substitute, 0.5 mM EDTA supplemented with PIC). Pull down eluates were separated by heating with NuPAGE™ LDS sample buffer (Invitrogen) + 4x DTT at 95°C for 15 min and analyzed by western blotting.

REFERENCES AND NOTES

1. R. Prakash, Y. Zhang, W. Feng, M. Jasin, Homologous recombination and human health: The roles of BRCA1, BRCA2, and associated proteins. *Cold Spring Harb. Perspect. Biol.* **7**, a016600 (2015). doi: 10.1101/cshperspect.a016600; pmid: 25833843
2. O. Belan *et al.*, Visualization of direct and diffusion-assisted RAD51 nucleation by full-length human BRCA2 protein. *Mol. Cell* **83**, 2925–2940.e8 (2023). doi: 10.1016/j.molcel.2023.06.031; pmid: 37499663

3. J. C. Bell, C. C. Dombrowski, J. L. Plank, R. B. Jensen, S. C. Kowalczykowski, BRCA2 chaperones RAD51 to single molecules of RPA-coated ssDNA. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2221971120 (2023). doi: 10.1073/pnas.2221971120; pmid: 36976771
4. A. A. Davies *et al.*, Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol. Cell* **7**, 273–282 (2001). doi: 10.1016/S1097-2765(01)00175-7; pmid: 11239456
5. R. B. Jensen, A. Carreira, S. C. Kowalczykowski, Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678–683 (2010). doi: 10.1038/nature09399; pmid: 20729832
6. J. Liu, T. Doty, B. Gibson, W. D. Heyer, Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nat. Struct. Mol. Biol.* **17**, 1260–1262 (2010). doi: 10.1038/nsmb.1904; pmid: 20729859
7. L. Y. Marmorstein, T. Ouchi, S. A. Aaronson, The BRCA2 gene product functionally interacts with p53 and RAD51. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13869–13874 (1998). doi: 10.1073/pnas.95.23.13869; pmid: 9811893
8. M. E. Moynahan, A. J. Pierce, M. Jasin, BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol. Cell* **7**, 263–272 (2001). doi: 10.1016/S1097-2765(01)00174-5; pmid: 11239455
9. S. K. Sharan *et al.*, Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature* **386**, 804–810 (1997). doi: 10.1038/386804a0; pmid: 9126738
10. F. Xia *et al.*, Deficiency of human BRCA2 leads to impaired homologous recombination but maintains normal nonhomologous end joining. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8644–8649 (2001). doi: 10.1073/pnas.151253498; pmid: 11447276
11. V. P. Yu *et al.*, Gross chromosomal rearrangements and genetic exchange between nonhomologous chromosomes following BRCA2 inactivation. *Genes Dev.* **14**, 1400–1406 (2000). doi: 10.1101/gad.14.11.1400; pmid: 10837032
12. M. Castroviejo-Bermejo *et al.*, A RAD51 assay feasible in routine tumor samples calls PARP inhibitor response beyond BRCA mutation. *EMBO Mol. Med.* **10**, e9172 (2018). doi: 10.15252/emmm.201809172; pmid: 30377213
13. C. Cruz *et al.*, RAD51 foci as a functional biomarker of homologous recombination repair and PARP inhibitor resistance in germline BRCA-mutated breast cancer. *Ann. Oncol.* **29**, 1203–1210 (2018). doi: 10.1093/annonc/mdy099; pmid: 29635390
14. K. A. Naipal *et al.*, Functional ex vivo assay to select homologous recombination-deficient breast tumors for PARP inhibitor treatment. *Clin. Cancer Res.* **20**, 4816–4826 (2014). doi: 10.1158/1078-0432.CCR-14-0571; pmid: 24963051
15. S. S. Yuan *et al.*, BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex in vivo. *Cancer Res.* **59**, 3547–3551 (1999). pmid: 10446958
16. H. E. Bryant *et al.*, Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913–917 (2005). doi: 10.1038/nature03443; pmid: 15829966
17. H. Farmer *et al.*, Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005). doi: 10.1038/nature03445; pmid: 15829967
18. E. Harvey-Jones *et al.*, Longitudinal profiling identifies co-occurring BRCA1/2 reversions, TP53BP1, RIF1 and PAXIP1 mutations in PARP inhibitor-resistant advanced breast cancer. *Ann. Oncol.* **35**, 364–380 (2024). doi: 10.1016/j.annonc.2024.01.003; pmid: 38244928
19. A. Ray Chaudhuri *et al.*, Replication fork stability confers chemoresistance in BRCA-deficient cells. *Nature* **535**, 382–387 (2016). doi: 10.1038/nature18325; pmid: 27443740
20. E. Duro *et al.*, Identification of the MMS22L-TONSL complex that promotes homologous recombination. *Mol. Cell* **40**, 632–644 (2010). doi: 10.1016/j.molcel.2010.10.023; pmid: 21055984
21. L. O'Donnell *et al.*, The MMS22L-TONSL complex mediates recovery from replication stress and homologous recombination. *Mol. Cell* **40**, 619–631 (2010). doi: 10.1016/j.molcel.2010.10.024; pmid: 21055983
22. W. Piwko *et al.*, The MMS22L–TONSL heterodimer directly promotes RAD51-dependent recombination upon replication stress. *The EMBO Journal* **35**, 2584–2601 (2016). doi: 10.1016/j.molcel.2018.01.031; pmid: 29478807
23. T. H. Huang *et al.*, The Histone Chaperones ASF1 and CAF-1 Promote MMS22L-TONSL-Mediated Rad51 Loading onto ssDNA during Homologous Recombination in Human Cells. *Mol. Cell* **69**, 879–892.e5 (2018). doi: 10.1016/j.molcel.2018.01.031; pmid: 29478807
24. G. Saredi *et al.*, H4K20me0 marks post-replicative chromatin and recruits the TONSL–MMS22L DNA repair complex. *Nature* **534**, 714–718 (2016). doi: 10.1038/nature18312; pmid: 27338793
25. J. Flygare *et al.*, Effects of HsRad51 overexpression on cell proliferation, cell cycle progression, and apoptosis. *Exp. Cell Res.* **268**, 61–69 (2001). doi: 10.1006/excr.2001.5265; pmid: 11461118
26. C. Richardson, J. M. Stark, M. Ommundsen, M. Jasin, Rad51 overexpression promotes alternative double-strand break repair pathways and genome instability. *Oncogene* **23**, 546–553 (2004). doi: 10.1038/sj.onc.1207098; pmid: 14724582
27. L. J. Barber *et al.*, RTEL1 maintains genomic stability by suppressing homologous recombination. *Cell* **135**, 261–271 (2008). doi: 10.1016/j.cell.2008.08.016; pmid: 18957201
28. Y. Hu *et al.*, RECQL5/Recql5 helicase regulates homologous recombination and suppresses tumor formation via disruption of Rad51 presynaptic filaments. *Genes Dev.* **21**, 3073–3084 (2007). doi: 10.1101/gad.1609107; pmid: 18003859
29. G.-L. Moldovan *et al.*, Inhibition of homologous recombination by the PCNA-interacting protein PARI. *Mol. Cell* **45**, 75–86 (2012). doi: 10.1016/j.molcel.2011.11.010; pmid: 22153967
30. D. S. Patel, S. M. Misenko, J. Her, S. F. Bunting, BLM helicase regulates DNA repair by counteracting RAD51 loading at DNA double-strand break sites. *J. Cell Biol.* **216**, 3521–3534 (2017). doi: 10.1083/jcb.201703144; pmid: 28912125
31. J. Simandlova *et al.*, FBH1 helicase disrupts RAD51 filaments in vitro and modulates homologous recombination in mammalian cells. *J. Biol. Chem.* **288**, 34168–34180 (2013). doi: 10.1074/jbc.M113.484493; pmid: 24108124
32. J. A. Sommers *et al.*, FANCJ uses its motor ATPase to destabilize protein-DNA complexes, unwind triplexes, and inhibit RAD51 strand exchange. *J. Biol. Chem.* **284**, 7505–7517 (2009). doi: 10.1074/jbc.M809019200; pmid: 19150983
33. K. Matsuzaki, A. Shinohara, M. Shinohara, Human AAA+ ATPase FIGNL1 suppresses RAD51-mediated ultra-fine bridge formation. *Nucleic Acids Res.* **52**, 5774–5791 (2024). doi: 10.1093/nar/gkae263; pmid: 38597669
34. K. Matsuzaki, S. Kondo, T. Ishikawa, A. Shinohara, Human RAD51 paralogue SWSAP1 fosters RAD51 filament by regulating the anti-recombinase FIGNL1 AAA+ ATPase. *Nat. Commun.* **10**, 1407 (2019). doi: 10.1038/s41467-019-09190-1; pmid: 30926776
35. E. Pinedo-Carpio *et al.*, FIRRM cooperates with FIGNL1 to promote RAD51 disassembly during DNA repair. *Sci. Adv.* **9**, eadf4082 (2023). doi: 10.1126/sciadv.adf4082; pmid: 37556550
36. C. Stok *et al.*, FIRRM/C1orf112 is synthetic lethal with PICH and mediates RAD51 dynamics. *Cell Rep.* **42**, 112668 (2023). doi: 10.1016/j.celrep.2023.112668; pmid: 37347663
37. J. D. Tischler *et al.*, FLIP(C1orf112)-FIGNL1 complex regulates RAD51 chromatin association to promote viability after replication stress. *Nat. Commun.* **15**, 866 (2024). doi: 10.1038/s41467-024-45139-9; pmid: 38286805
38. J. Yuan, J. Chen, FIGNL1-containing protein complex is required for efficient homologous recombination repair. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 10640–10645 (2013). doi: 10.1073/pnas.1220662110; pmid: 23754376
39. Z. Zhou *et al.*, C1orf112 teams up with FIGNL1 to facilitate RAD51 filament disassembly and DNA interstrand cross-link repair. *Cell Rep.* **42**, 112907 (2023). doi: 10.1016/j.celrep.2023.112907; pmid: 37515771
40. M. Ito *et al.*, FIGNL1 AAA+ ATPase remodels RAD51 and DMC1 filaments in pre-meiotic DNA replication and meiotic recombination. *Nat. Commun.* **14**, 6857 (2023). doi: 10.1038/s41467-023-42576-w; pmid: 37891173
41. A. Zainu *et al.*, FIGNL1-FIRRM is essential for meiotic recombination and prevents DNA damage-independent RAD51 and DMC1 loading. *Nat. Commun.* **15**, 7015 (2024). doi: 10.1038/s41467-024-51458-8; pmid: 39147779
42. R. Kumar, M. Duhamel, E. Coutant, E. Ben-Nahia, R. Mercier, Antagonism between BRCA2 and FIGL1 regulates homologous recombination. *Nucleic Acids Res.* **47**, 5170–5180 (2019). doi: 10.1093/nar/gkz225; pmid: 30941419
43. A. Carver *et al.*, Molecular basis of FIGNL1 in dissociating RAD51 from DNA and chromatin. *Science* **387**, 426–431 (2025). doi: 10.1126/science.adr7920; pmid: 39636933
44. A. M. Heijink *et al.*, Sister chromatid exchanges induced by perturbed replication can form independently of BRCA1, BRCA2 and RAD51. *Nat. Commun.* **13**, 6722 (2022). doi: 10.1038/s41467-022-34519-8; pmid: 36344511
45. B. Evers *et al.*, A high-throughput pharmaceutical screen identifies compounds with specific toxicity against BRCA2-deficient tumors. *Clin. Cancer Res.* **16**, 99–108 (2010). doi: 10.1158/1078-0432.CCR-09-2434; pmid: 20008842
46. C. S. Y. Lo, N. Taneja, A. Ray Chaudhuri, Enhancing quantitative imaging to study DNA damage response: A guide to automated liquid handling and imaging. *DNA Repair* **144**, 103769 (2024). doi: 10.1016/j.dnarep.2024.103769; pmid: 39395383
47. L. I. Toledo *et al.*, ATR prohibits replication catastrophe by preventing global exhaustion of RPA. *Cell* **155**, 1088–1103 (2013). doi: 10.1016/j.cell.2013.10.043; pmid: 24267891
48. A. Canela *et al.*, DNA Breaks and End Resection Measured Genome-wide by End Sequencing. *Mol. Cell* **63**, 898–911 (2016). doi: 10.1016/j.molcel.2016.06.034; pmid: 27477910
49. J. S. Iacovoni *et al.*, High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. *EMBO J.* **29**, 1446–1457 (2010). doi: 10.1038/emboj.2010.38; pmid: 20360682
50. S. Raynard, P. Sung, Assay for human Rad51-mediated DNA displacement loop formation. *Cold Spring Harb Protoc.* **4**, pdb.prot5120 (2009). doi: 10.1038/emboj.2010.38; pmid: 20360682
51. P. Chi, S. Van Komen, M. G. Sehorn, S. Sigurdsson, P. Sung, Roles of ATP binding and ATP hydrolysis in human Rad51 recombinase function. *DNA Repair* **5**, 381–391 (2006). doi: 10.1016/j.dnarep.2005.11.005; pmid: 16388992
52. J. M. Stark *et al.*, ATP hydrolysis by mammalian RAD51 has a key role during homology-directed DNA repair. *J. Biol. Chem.* **277**, 20185–20194 (2002). doi: 10.1074/jbc.M112132200; pmid: 11923292
53. N. Siaud *et al.*, Plasticity of BRCA2 function in homologous recombination: Genetic interactions of the PALB2 and DNA binding domains. *PLOS Genet.* **7**, e1002409 (2011). doi: 10.1371/journal.pgen.1002409; pmid: 22194698
54. Y. Kwon *et al.*, DNA binding and RAD51 engagement by the BRCA2 C-terminus orchestrate DNA repair and replication fork preservation. *Nat. Commun.* **14**, 432 (2023). doi: 10.1038/s41467-023-36211-x; pmid: 36702902

55. T. Thorslund *et al.*, The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. *Nat. Struct. Mol. Biol.* **17**, 1263–1265 (2010). doi: 10.1038/nsmb.1905; pmid: 20729858
56. K. Nakanishi, F. Cavallo, E. Brunet, M. Jasin, in *DNA Recombination: Methods and Protocols*, H. Tsubouchi, Ed. (Humana Press, 2011), pp. 283–291.doi: 10.1007/978-1-61779-129-1_16
57. A. Suzuki *et al.*, Brca2 is required for embryonic cellular proliferation in the mouse. *Genes Dev.* **11**, 1242–1252 (1997). doi: 10.1101/gad.11.10.1242; pmid: 9171369
58. S. G. Kuznetsov, P. Liu, S. K. Sharan, Mouse embryonic stem cell-based functional assay to evaluate mutations in BRCA2. *Nat. Med.* **14**, 875–881 (2008). doi: 10.1038/nm.1719; pmid: 18607349
59. K. Biswas *et al.*, A comprehensive functional characterization of BRCA2 variants associated with Fanconi anemia using mouse ES cell-based assay. *Blood* **118**, 2430–2442 (2011). doi: 10.1182/blood-2010-12-324541; pmid: 21719596
60. E. W. Schmid, J. C. Walter, Predictomes, a classifier-curated database of AlphaFold-modeled protein-protein interactions. *Mol. Cell* **85**, 1216–1232.e5 (2025). doi: 10.1016/j.molcel.2025.01.034; pmid: 40015271
61. M. Reuter *et al.*, BRCA2 diffuses as oligomeric clusters with RAD51 and changes mobility after DNA damage in live cells. *J. Cell Biol.* **207**, 599–613 (2014). doi: 10.1083/jcb.201405014; pmid: 25488918
62. S. Halder *et al.*, Double-stranded DNA binding function of RAD51 in DNA protection and its regulation by BRCA2. *Mol. Cell* **82**, 3553–3565.e5 (2022). doi: 10.1016/j.molcel.2022.08.014; pmid: 36070766
63. M. A. Longo *et al.*, BRCA2 C-terminal clamp restructures RAD51 dimers to bind B-DNA for replication fork stability. *Mol. Cell* **85**, 2080–2096.e6 (2025). doi: 10.1016/j.molcel.2025.05.010; pmid: 40441151
64. S. Mijic *et al.*, Replication fork reversal triggers fork degradation in BRCA2-defective cells. *Nat. Commun.* **8**, 859 (2017). doi: 10.1038/s41467-017-01164-5; pmid: 29038466
65. K. Schlacher *et al.*, Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell* **145**, 529–542 (2011). doi: 10.1016/j.cell.2011.03.041; pmid: 21565612
66. L. A. Greenhough *et al.*, Structure and function of the RAD51B-RAD51C-RAD51D-XRCC2 tumour suppressor. *Nature* **619**, 650–657 (2023). doi: 10.1038/s41586-023-06179-1; pmid: 37344587
67. M. A. Longo *et al.*, RAD51C-XRCC3 structure and cancer patient mutations define DNA replication roles. *Nat. Commun.* **14**, 4445 (2023). doi: 10.1038/s41467-023-40096-1; pmid: 37488098
68. Y. Rawal *et al.*, Structural insights into BCDX2 complex function in homologous recombination. *Nature* **619**, 640–649 (2023). doi: 10.1038/s41586-023-06219-w; pmid: 37344589
69. T. G. Meijer *et al.*, Direct Ex Vivo Observation of Homologous Recombination Defect Reversal After DNA-Damaging Chemotherapy in Patients With Metastatic Breast Cancer. *JCO Precis. Oncol.* **3**, 1–12 (2019). doi: 10.1200/PO.18.00268; pmid: 35100677
70. J. Bhin *et al.*, Multi-omics analysis reveals distinct non-reversion mechanisms of PARPi resistance in BRCA1- versus BRCA2-deficient mammary tumors. *Cell Rep.* **42**, 112538 (2023). doi: 10.1016/j.celrep.2023.112538; pmid: 37209095
71. J. Jonkers *et al.*, Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nat. Genet.* **29**, 418–425 (2001). doi: 10.1038/ng747; pmid: 11694875
72. K. Nakanishi, F. Cavallo, E. Brunet, M. Jasin, Homologous recombination assay for interstrand cross-link repair. *Methods Mol. Biol.* **745**, 283–291 (2011). doi: 10.1007/978-1-61779-129-1_16; pmid: 21660700
73. I. Ceppi *et al.*, Mechanism of BRCA1-BARD1 function in DNA end resection and DNA protection. *Nature* **634**, 492–500 (2024). doi: 10.1038/s41586-024-07909-9; pmid: 39261728
74. D. R. Stirling *et al.*, CellProfiler 4: Improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 433 (2021). doi: 10.1186/s12859-021-04344-9; pmid: 34507520
75. J. Abramson *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024). doi: 10.1038/s41586-024-07487-w; pmid: 38718835
76. G. Kim *et al.*, Easy and accurate protein structure prediction using ColabFold. *Nat. Protoc.* **20**, 620–642 (2025). doi: 10.1038/s41596-024-01060-5; pmid: 39402428
77. L. C. Xue, J. P. Rodrigues, P. L. Kastritis, A. M. Bonvin, A. Vangone, PRODIGY: A web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics* **32**, 3676–3678 (2016). doi: 10.1093/bioinformatics/btw514; pmid: 27503228
78. C. Elfmann, J. Stülke, PAE viewer: A webserver for the interactive visualization of the predicted aligned error for multimer structure predictions and crosslinks. *Nucleic Acids Res.* **51**, W404–W410 (2023). doi: 10.1093/nar/gkad350; pmid: 37140053
79. C. Pinto, K. Kasaciunaite, R. Seidel, P. Cejka, Human DNA2 possesses a cryptic DNA unwinding activity that functionally integrates with BLM or WRN helicases. *eLife* **5**, e18574 (2016). doi: 10.7554/eLife.18574; pmid: 27612385
80. G. Reginato *et al.*, HLTF disrupts Cas9-DNA post-cleavage complexes to allow DNA break processing. *Nat. Commun.* **15**, 5789 (2024). doi: 10.1038/s41467-024-50080-y; pmid: 38987539
81. A. V. Mazin, E. Zaitseva, P. Sung, S. C. Kowalczykowski, Tailed duplex DNA is the preferred substrate for Rad51 protein-mediated homologous pairing. *EMBO J.* **19**, 1148–1156 (2000). doi: 10.1093/emboj/19.5.1148; pmid: 10698955
82. C. Mukherjee *et al.*, RIF1 promotes replication fork protection and efficient restart to maintain genome stability. *Nat. Commun.* **10**, 3287 (2019). doi: 10.1038/s41467-019-11246-1; pmid: 31337767
83. A. R. Chaudhuri, R. Kuthethur, FIGNL1 inhibits homologous recombination in BRCA2 deficient cells by dissociating RAD51 filaments. Zenodo (2025); https://zenodo.org/records/15462520.
84. A. R. Chaudhuri, K. M. Poluri, R. Kuthethur, N. Nagar, FIGNL1 inhibits homologous recombination in BRCA2 deficient cells by dissociating RAD51 filaments (Version 1). Zenodo (2025); https://zenodo.org/records/15454978

ACKNOWLEDGMENTS

The authors thank Optical Imaging Centre of the Erasmus MC for assistance and microscope access for the study. The authors thank M. Jasin for the miBRCA2 expression construct, I. Ceppi and M.R.D. Sritto for cloning the miBRCA2 construct, and W. Vermeulen for critical reading of the manuscript and discussions. **Funding:** This work was funded by the following: NWO (Dutch research Council) VIDI award (Vidi.193.131) (to A.R.C.); Ammodo Science Award (to A.R.C., R.Ka., and J.H.G.L.); and Josephine Nefkens Cancer Program (to A.R.C., R.Ka., and J.H.G.L.); R.Ka. is also funded by the Oncode Institute, which is partly financed by the Dutch Cancer Society (to K.W.F.); Canadian Institutes of Health Research Grants (PDT_180310) to A.F.T. A.F.T. holds a Canada Research Chair in Molecular Virology and Genomic Instability (Tier 2) and the Foundation J.-Louis Lévesque also supports her lab work. This work was also funded by SERB-INDIA (Science and Engineering Research Board) research grants STR/2022/000008 and CRG/2022/003028 to KMP; research in the PC laboratory is funded by the Swiss National Science Foundation (SNSF) (grants 310030_207588 and 310030_205199) and the European Research Council (ERC) (grant 101018257). The S.K. Sharan lab is funded by Intramural Research Program, Center for Cancer Research, National Cancer Institute, US National Institutes of Health. **Author contributions:** Conceptualization: A.R.C. and R.Ku. Methodology: R.Ku., A.A., S.K.Se., C.F., N.N., S.B., and J.D. Investigation: R.Ku., A.A., S.K.Se., C.F., N.N., O.I., S.N., S.B., E.-M.M., K.dK., A.T., and X.W. Funding acquisition: A.R.C., P.C., S.K.Sha., K.M.P., J.H.G.L., R.Ka., and A.F.T. Project administration A.R.C. Supervision: A.R.C., P.C., S.K.Sha., K.M.P., J.H.G.L., and R.Ka. Writing-original draft: A.R.C. with contributions from R.Ku. and C.F. Writing-review & editing: P.C., S.K.Sha., K.M.P., J.H.G.L., and R.Ka. and all other authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** CellProfiler pipeline templates used for foci detection and colocalization analysis are deposited in the zenodo repository (*83*). All the relevant materials used for AlphaFold structural data analysis are deposited in the Zenodo repository (*84*). Cell lines and plasmids created in this study are available upon request from A.R.C. under materials transfer agreement with Erasmus University Medical Centre, Rotterdam, NL. Transfer of cell lines and plasmids may be subject to third party obligations. All other data are available in the main text or the supplementary materials. **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US gover nMent works. https://www.science.org/content/page/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adt1210
Figs. S1 to S18; Tables S1 to S6; MDAR Reproducibility Checklist
Submitted 13 September 2024; accepted 24 July 2025
10.1126/science.adt1210

RESEARCH ARTICLE SUMMARY

PROTEIN CHAPERONES

Structural dissection of αβ-tubulin heterodimer assembly and disassembly by human tubulin–specific chaperones

Yeonjae Seong†, Hyunmin Kim†, Kyumi Byun, Yeon-Woo Park, Soung-Hun Roh*

Full article and list of author affiliations: https://doi.org/10.1126/science.ady2708

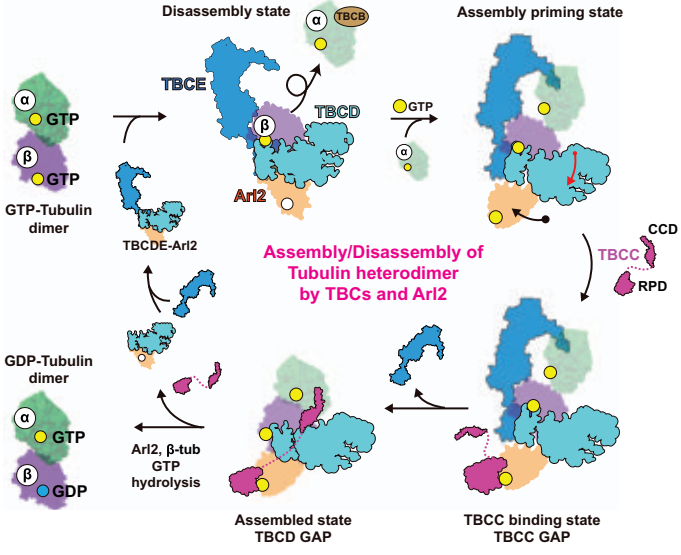
INTRODUCTION: Microtubules are essential cytoskeletal polymers critical for cell division, intracellular transport, and maintenance of cell shape. They are composed of αβ-tubulin heterodimers, each consisting of an α-tubulin and a β-tubulin monomer. Assembly and disassembly of tubulin heterodimers are assisted by tubulin binding cofactors (TBCs): TBCA, TBCB, TBCC, TBCD, and TBCE. In addition, a small guanosine triphosphatase (GTPase), Arl2, is involved in this pathway.

RATIONALE: TBCs can mediate both guanosine triphosphate (GTP)–independent disassembly and GTP-dependent assembly of αβ-tubulin heterodimers. Although this process is essential for maintaining cytoskeletal homeostasis, the mechanisms involved remain poorly understood. We used cryo–electron microscopy (cryo-EM) to analyze reconstituted human TBC complexes under defined nucleotide conditions and captured structural snapshots of the disassembly and assembly pathway.

RESULTS: We purified human TBCDE-Arl2 complex and TBCC from mammalian cells and expressed TBCA and TBCB individually in *Escherichia coli*. These components were reconstituted with αβ-tubulin under defined GTP conditions, yielding six cryo-EM structures with 2.1- to 2.5-Å resolution. In the absence of GTP, we captured State I (TBCD–Arl2–β-tub) and State II (TBCDE–Arl2–β-tub), in which only β-tubulin was resolved, indicating disassembly of the heterodimer. In these states, β-tubulin is stabilized through electrostatic complementarity, as its highly negatively charged surface engages positively charged regions on TBCD (Armadillo-repeat domains AD1 and AD3), which compensates for the loss of α-tubulin. These interactions remodel the S9-H10 motif of β-tubulin, forming a new interface that mitigates its intrinsic instability. Upon addition of GTP, Arl2 undergoes a major conformational change in State III (TBCDE–$Arl2_{GTP}$–β-tub), which displaces the AD2 domain of TBCD, enabling α-tubulin binding. This leads to State IV (TBCDE–$Arl2_{GTP}$–αβ-tub), in which α-tubulin is resolved, revealing a partially reassembled heterodimer. With the addition of TBCC and a post–GTP hydrolysis mimic (GTP-AlFx), we resolved State V (TBCDEC–$Arl2_{GDP\text{-}AlFx}$–αβ-tub) and State VI (TBCDC–$Arl2_{GDP\text{-}AlFx}$–αβ-tub). In State V, the regulatory protein domain (RPD) of TBCC binds the GTP-bound Arl2 and contributes to the catalytic site, whereas in State VI, the coiled-coil domain (CCD) acts as a molecular clamp bridging α- and β-tubulin and displacing TBCE. These transitions complete formation of the native αβ-tubulin interface and finalize heterodimer assembly. Lastly, our structural analysis revealed that TBCD functions as a GTPase-activating protein for β-tubulin. β-Tubulin GTP hydrolysis is coupled to Arl2's GTPase activity, establishing a coordinated checkpoint that ensures that only fully assembled and mature αβ-tubulin heterodimers are released.

CONCLUSION: Based on our structures, we propose a mechanism by which TBCs mediate both the GTP-independent disassembly and GTP-dependent reassembly of αβ-tubulin heterodimers. Arl2 acts as a central regulator that switches the complex between these opposing states, and the coordinated GTPase activities of Arl2 and β-tubulin provide quality control that ensures precise regulation of tubulin heterodimer homeostasis. □

*Corresponding author. Email: shroh@snu.ac.kr †These authors contributed equally to this work. Cite this article as Y. Seong *et al.*, *Science* **390**, eady2708 (2025). DOI: 10.1126/science.ady2708



Schematic overview of the αβ-tubulin heterodimer assembly-disassembly cycle. In the disassembly state, αβ-tubulin is dissociated and monomeric β-tubulin is stabilized by TBCDE-Arl2. GTP binding induces the assembly priming state, enabling α-tubulin recruitment. TBCC engagement completes the assembly of native αβ-tubulin heterodimer and facilitates the release of TBCE. Coupled GTP hydrolysis in Arl2, and β-tubulin ensures the release of a functionally competent αβ-tubulin heterodimer through a quality control checkpoint.

PROTEIN CHAPERONES

Structural dissection of αβ-tubulin heterodimer assembly and disassembly by human tubulin-specific chaperones

Yeonjae Seong[1]†, Hyunmin Kim[1]†‡, Kyumi Byun[1], Yeon-Woo Park[1], Soung-Hun Roh[1,2]*

Microtubule assembly requires a set of chaperones known as tubulin-binding cofactors (TBCs). We used cryo–electron microscopy to visualize how human TBCD, TBCE, TBCC, and guanosine triphosphatase (GTPase) Arl2 mediate αβ-tubulin assembly and disassembly. We captured multiple conformational states, revealing how TBCs orchestrate tubulin heterodimer biogenesis. TBCD stabilizes monomeric β-tubulin and scaffolds the other cofactors. Guanosine triphosphate (GTP) binding to Arl2 induces conformational changes that toggle the complex between assembly and disassembly. TBCD and TBCE guide α- and β-tubulin into a partially assembled interface, and TBCC, acting as a molecular clamp, completes the heterodimer. TBCD also functions as a GTPase activating protein for β-tubulin. β-tubulin GTP hydrolysis is coupled to Arl2's GTPase activity, establishing a checkpoint that ensures that only fully matured heterodimers proceed. These findings provide a structural framework for tubulin heterodimer biogenesis and recycling, supporting cytoskeletal proteostasis.

Microtubules are fundamental components of the cytoskeleton of eukaryotic cells. Not only do they maintain cell shape, but they allow the intracellular transport of vesicles and organelles attached to molecular motors and form the spindle that segregates chromosomes into daughter cells at cell division (*1*–*3*), a process that illustrates their dynamic nature, rapidly switching between growth and shrinkage (*4*, *5*).

Microtubules assemble from a cytoplasmic pool of tubulin heterodimers, each comprising an α- and β-tubulin monomer (*6*, *7*). Newly synthesized tubulin monomers require cytosolic chaperonin TCP1 ring complex (TRiC/CCT) for their folding (*8*–*10*), and they are highly unstable (*11*, *12*) until they are assembled into functional αβ-tubulin heterodimers by a set of specific chaperones called tubulin-binding cofactors (TBCs) (*13*). In humans, these cofactors include TBCA, TBCB, TBCC, TBCD, and TBCE, which have specialized roles exclusively in tubulin biogenesis (*13*). Additionally, the small guanosine triphosphatase (GTPase) adenosine diphosphate (ADP) ribosylation factor–like protein 2 (Arl2) is involved in regulating tubulin heterodimer biogenesis (*14*). However, Arl2 is not classified as a canonical TBC because it performs multiple cellular functions beyond tubulin biogenesis, including roles in mitochondria (*15*, *16*). Currently, we have a broad understanding of the functions of these proteins (Fig. 1A). TBCA and TBCB stabilize β- and α-tubulin monomers, respectively, upon their release from TRiC/CCT (*17*) and deliver them to TBCD and TBCE (*11*), which together assemble the monomers into a heterodimer. Then TBCC activates GTP hydrolysis by the complex (*11*), facilitating release of the functional, native αβ-tubulin heterodimer (*13*, *18*).

TBCs also mediate the disassembly and reassembly of existing αβ-tubulin heterodimers. For example, TBCB and TBCE can disassemble αβ-tubulin into monomers in a guanosine triphosphate (GTP)–independent manner (*19*), and in the presence of GTP, TBCC promotes heterodimer reassembly (*11*). During this process, GTP binding and hydrolysis occurs at Arl2, with Arl2 acting as a molecular switch through its interaction with TBCD (*20*). Although α- and β-tubulin both have GTP binding sites and β-tubulin has intrinsic GTPase activity (*21*), the function of GTP in αβ-tubulin heterodimer reassembly appears to involve Arl2, which is found in a complex with TBCD and β-tubulin in most cells and tissues (*22*). Hydrogen-deuterium exchange experiments coupled with mass spectrometry concluded that Arl2 exchanges GTP in a complex containing TBCD–Arl2–β-tubulin (*23*).

Biochemical studies have shown that TBCD exists in several distinct complexes, including TBCD–Arl2–β-tubulin (*22*), and TBCE forms complexes with TBCB and α-tubulin (*24*). Moreover, TBCD, Arl2, and TBCE can combine with TBCC and α- and β-tubulin to form what has been called a "supercomplex" with GTPase activity, which assembles αβ-tubulin heterodimers (*25*). In principle, such a supercomplex might also work in reverse to disassemble the heterodimers, ensuring homeostasis of the cytoplasmic pool.

A recent cryo–electron microscopy (cryo-EM) study has begun to provide structural snapshots of yeast TBC complexes (*26*), but a comprehensive mechanistic understanding of the human TBC system remains limited. To address this problem, we used a cryo-EM approach to analyze the process of disassembly and assembly of αβ-tubulin heterodimers by complexes containing purified, recombinant human TBCD, TBCE, Arl2, and TBCC in the absence and presence of GTP.

Reconstitution of six distinct TBC-tubulin complex structures

To obtain the human tubulin cofactors required to reconstitute tubulin heterodimer biogenesis (Fig. 1A), we purified a complex of TBCD, TBCE, and Arl2 (TBCDE-Arl2) and TBCC from the human cell line human embryonic kidney (HEK) 293S and expressed human TBCA and TBCB individually in *Escherichia coli* (fig. S1, A and B). Individual expression of TBCE and TBCD in HEK293S resulted in low yields; however, co-expression of TBCD, TBCE, and Arl2 as a complex markedly enhanced protein production and stability, as observed previously for the yeast cofactors (*14*). We then reconstituted the TBCDE-Arl2 complex with αβ-tubulin in the absence and presence of TBCA, TBCB, TBCC, and GTP (Fig. 1B) and used cryo-EM to obtain high-resolution reconstructions (2.1 to 2.5 Å) corresponding to six distinct complex structures (States I to VI) containing either β-tubulin monomers (States I to III) or αβ-tubulin heterodimers (States IV to VI) bound to the TBCs (Fig. 1C, figs. S2 to S4, and movie S1).

In the absence of GTP, reconstitution of TBCDE-Arl2 with αβ-tubulin yielded two structures, State I (TBCD–Arl2–β-tub) and State II (TBCDE–Arl2–β-tub), in which only β-tubulin was resolved, suggesting dissociation of the heterodimer. The presence of GTP produced State III (TBCDE–$Arl2_{GTP}$–β-tub), in which only β-tubulin was resolved, and State IV (TBCDE–$Arl2_{GTP}$–αβ-tub), in which both α- and β-tubulin were clearly resolved. In State IV, β-tubulin was anchored to TBCD, as observed in States I to III, and α-tubulin interacted primarily with TBCE. In both State III and State IV, there were major conformational changes in Arl2 due to GTP binding. Because TBCC drives rapid GTP hydrolysis by Arl2, we used GTP with aluminum fluoride (AlFx) condition in the TBCDE-Arl2 complex with TBCC to mimic a posthydrolysis-like state (*14*), thus minimizing structural heterogeneity and allowing us to determine a high-resolution structure. Addition of TBCC to TBCDE–$Arl2_{GDP\text{-}AlFx}$–αβ-tub (GDP, guanosine diphosphate) resulted in State V, which contained TBCC and TBCE (TBCDEC–$Arl2_{GDP\text{-}AlFx}$–αβ-tub), and State VI, which lacked TBCE (TBCDC–$Arl2_{GDP\text{-}AlFx}$–αβ-tub). In State V,

[1]School of Biological Sciences, Seoul National University, Seoul, Republic of Korea. [2]Institute of Molecular Biology and Genetics, Seoul National University, Seoul, Republic of Korea.
*Corresponding author. Email: shroh@snu.ac.kr †These authors contributed equally to this work. ‡Present address: Molecular Systems Biology Unit, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany.

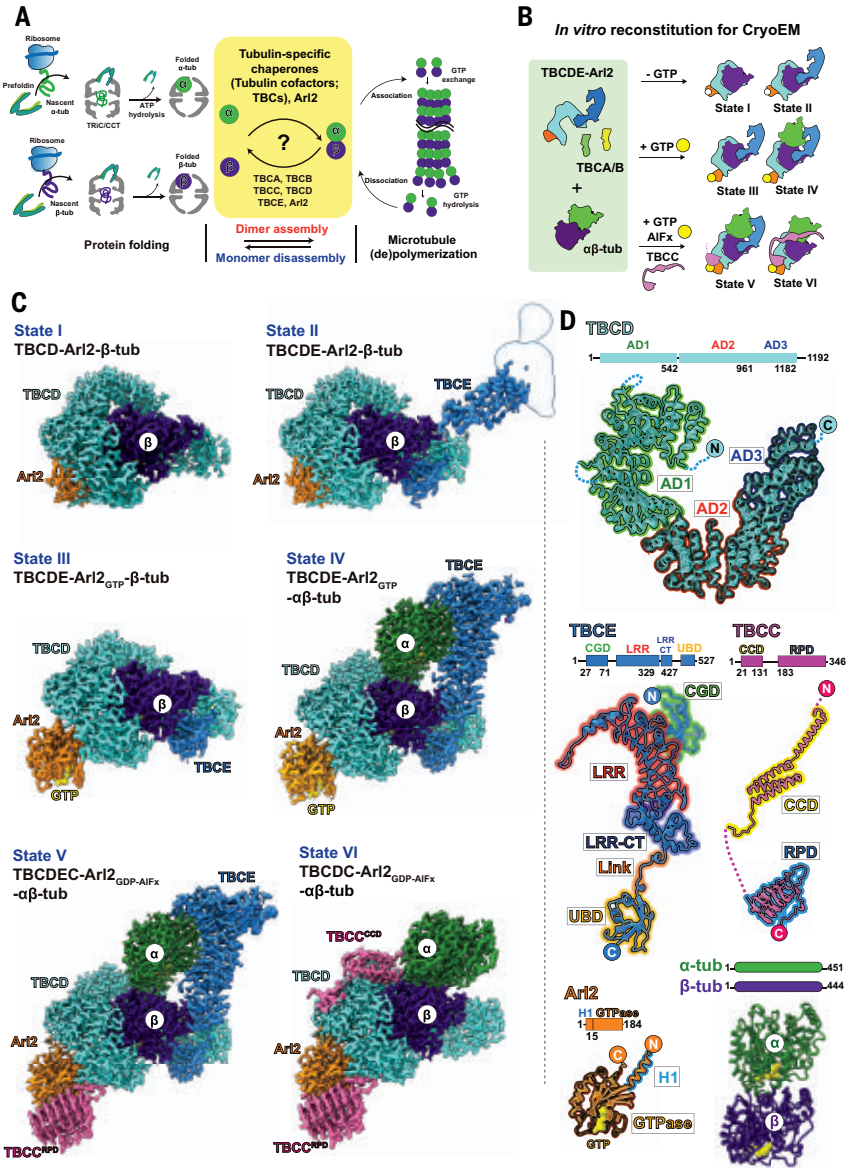


Fig. 1. Reconstitution of six distinct TBC-tubulin complex structures. (**A**) Tubulin biogenesis, from chaperonin-assisted monomer folding to αβ-tubulin heterodimer assembly and eventual microtubule polymerization. After folding, tubulin monomers undergo several steps, guided by tubulin-binding cofactors (TBCs), to form the heterodimer. (**B**) Experimental scheme for in vitro reconstitution of tubulin biogenesis using TBCs and αβ-tubulin for cryo-EM analyses. (**C**) Cryo-EM structures reveal six distinct states of tubulin heterodimer assembly and disassembly, each orchestrated by complex of TBCD, TBCE, TBCC, and Arl2. The color code is as follows: TBCD (cyan), TBCE (blue), Arl2 (orange), TBCC (pink), α-tubulin (green), and β-tubulin (purple). The same color scheme is used throughout the figures. (**D**) Domain architecture of TBCs and the αβ-tubulin heterodimer. TBCD is composed of three Armadillo-repeat domains (AD1 to AD3). TBCE comprises an N-terminal Cap-Gly domain, a horseshoe-shaped LRR domain, and a C-terminal UBD. Arl2 contains an N-terminal helix (H1) and a GTPase domain. TBCC has an N-terminal CCD and a C-terminal RPD, connected by a flexible linker.

only the regulatory domain of TBCC was resolved and could be seen interacting with Arl2. In State VI, however, the coiled-coil domain of TBCC was resolved; this domain could be seen clamping both α- and β-tubulin into a native heterodimer–like conformation, with TBCE no longer present in the complex.

In these six structural states, each TBC has a specific architecture reflecting its function in tubulin heterodimer assembly (Fig. 1D). TBCD, with its three curved Armadillo-repeat domains (AD1 to AD3), serves as the central scaffold, providing a binding platform for TBCE, TBCC, Arl2, and β-tubulin. TBCE, with its horseshoe-shaped structure, comprises an N-terminal Cap-Gly domain, an extended leucine-rich repeat (LRR) that interacts with both tubulin monomers, and a C-terminal ubiquitin-like domain (UBD) that binds to TBCD. Arl2 features a protruding N-terminal α helix (H1, residues 1 to 14) and an ADP ribosylation factor–like domain that contains its GTP binding, catalytic site (*27*). Lastly, TBCC consists of an N-terminal coiled-coil domain (CCD), a flexible linker region, and a C-terminal regulatory protein domain (RPD).

TBC complexes disassemble tubulin heterodimers and stabilize β-tubulin monomers

Previous studies have shown that both TBCD and TBCE can dissociate αβ-tubulin independently of GTP (*11*, *28*). Consistent with these findings, our native polyacrylamide gel electrophoresis (PAGE) analysis revealed that reconstitution of TBCDE-Arl2 with αβ-tubulin led to heterodimer dissociation and exclusive binding of TBC complex to β-tubulin in the absence of GTP; no association with α-tubulin was detected (Fig. 2A). This incubation results in formation of distinct TBCD–Arl2–β-tub and TBCDE–Arl2–β-tub complexes, further confirmed by mass spectrometry and biochemical analysis (Fig. 2A and fig. S5, A to C). Moreover, addition of TBCA or TBCB enhanced this phenomenon, decreasing the remaining amount of αβ-tubulin heterodimer and resulting in formation of TBCB–α-tub (Fig. 2A and fig. S5B). These data are consistent with observations that TBCE dissociates αβ-tubulin more effectively when TBCB is present (*24*).

The State I (TBCD–Arl2–β-tub) and State II (TBCDE–Arl2–β-tub) models capture structures in which the αβ-tubulin has been disassembled and β-tubulin is bound to TBCs in the absence of GTP (Fig. 2, B and C). We also observed two-dimensional (2D) class averages resembling tubulin monomers, likely representing dissociated α-tubulin (fig. S5D). Owing to the few orientations and the size constraints for cryo-EM reconstruction, however, we were unable to generate a 3D map of these monomeric forms. In States I and II, TBCD engages β-tubulin primarily through its AD1 and AD3 domains; AD2 does not interact with β-tubulin.

AD1 positions β-tubulin and Arl2 on opposite sides of the Armadillo domain and brings them into close proximity. State II has a similar overall architecture to State I, but TBCE is clearly resolved and associates with the AD3 domain of TBCD by means of its UBD. The LRR and Cap-Gly domains of TBCE extend outward through a flexible linker, adopting a structurally dynamic conformation. The interactions between the TBCs and β-tubulin are driven primarily by interactions between specific residues (Fig. 2C) and overall electrostatic complementarity (Fig. 2D). β-tubulin has a highly negative surface charge, whereas the AD1 and AD3 domains of TBCD have positively charged surfaces, generating multiple salt bridges and polar contacts that contribute to the strong binding affinity between TBCD and β-tubulin. The UBD of TBCE directly contacts β-tubulin through several interactions, including multiple salt bridges. Moreover, AD3 has a distinct negatively charged surface patch, which serves as a key

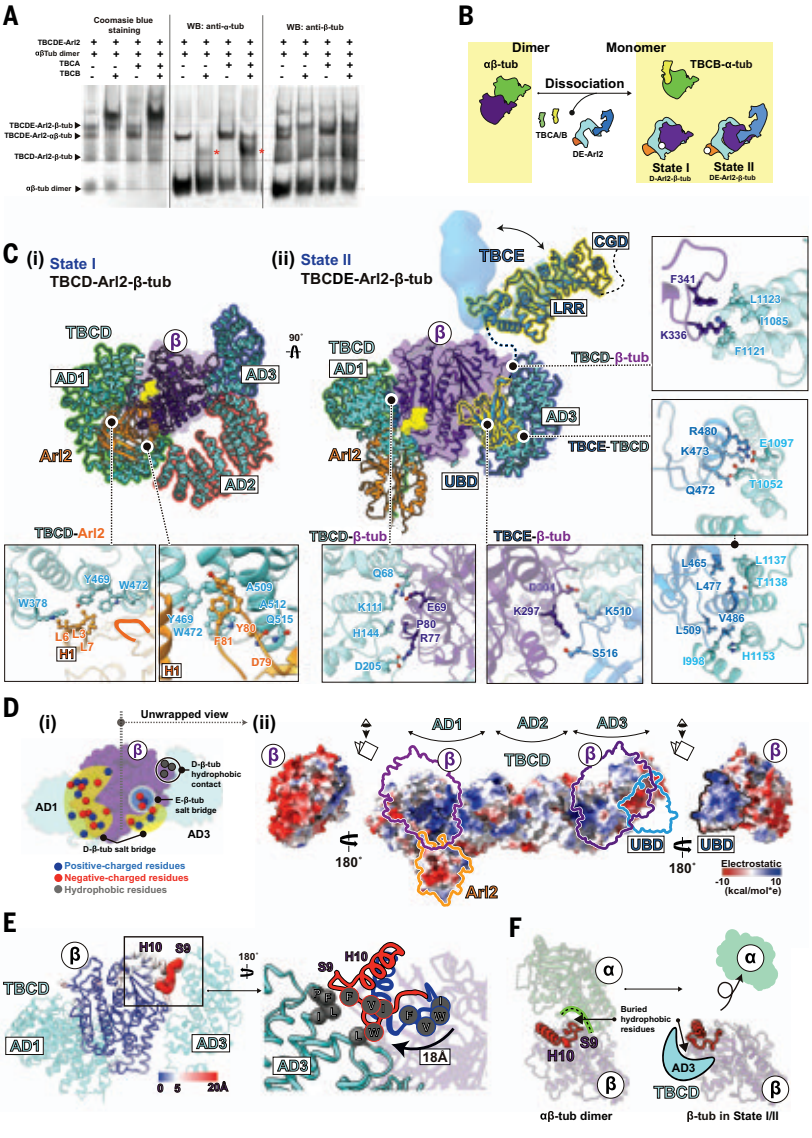


Fig. 2. TBC complexes disassemble tubulin heterodimers and stabilize β-tubulin monomers. (**A**) Native PAGE and western blot analysis of tubulin heterodimer dissociation upon incubation with TBCDE-Arl2. When TBCA, TBCB, or both are also present, the dissociation of αβ-tubulin is further facilitated, yielding complexes such as TBCDE–Arl2–β-tub, TBCDE–Arl2–αβ-tub, and TBCD–Arl2–β-tub. The asterisk indicates the formation of TBCB–α-tub. (**B**) Experimental scheme for reconstituting TBC-DE–Arl2 and αβ-tubulin to capture dissociation states by cryo-EM. (**C**) Cryo-EM structures of two dissociated tubulin states: (i) State I (TBCD–Arl2–β-tub) and (ii) State II (TBCDE–Arl2–β-tub). TBCE LRR displays dynamic features. Boxes show zoomed-in views of detailed interactions between TBCD and the other components or TBCE and β-tubulin. Black dashed lines with rounded ends indicate the locations of the interactions within the complex. (**D**) (i) Schematic representation of key interaction sites between TBCs and β-tubulin. The dashed line denotes the radial axis upon which the molecular surface was projected to generate the unwrapped surface view in (ii). (ii) Unwrapped electrostatic surface representation of the TBCDE–Arl2–β-tub. Open-book views for AD1 and AD3 are displayed to show the surface charge complementarity between TBCD and other components. The footprint of β-tubulin is contoured in purple; Arl2, in orange; and the UBD of TBCE, in blue. Arrows indicate the overall curvature of AD1 to AD3 in TBCD. (**E**) Remodeling and stabilization of β-tubulin by TBCD. The S9-H10 motif of β-tubulin flips outward and is partially stabilized by AD3 of TBCD through hydrophobic interactions. (Left) The S9-H10 motif is colored by RMSD between β-tubulin in TBCs. (Right) Hydrophobic residues in the motif and the RMSD relative to the free heterodimer conformation are shown. (**F**) Schematic of tubulin heterodimer dissociation mediated by the TBCDE-Arl2. Upon binding to the TBC complex, α-tubulin is dissociated, and the S9-H10 motif of β-tubulin is reconfigured and stabilized by TBCD. Single-letter abbreviations for the amino acid residues referenced throughout the paper and figures are as follows: D, Asp; A, Ala; E, Glu; Q, Gln; L, Leu; S, Ser; H, His; R, Arg; K, Lys; F, Phe; V, Val; T, Thr; W, Trp; Y, Tyr; P, Pro.

docking site for UBD. This complementary attraction stabilizes the complex, ensuring the specificity and strength of interaction between the TBCs and β-tubulin.

Previous biochemical studies have indicated that monomeric tubulin is inherently unstable (*11*, *12*); thus, we investigated how TBCs stabilize monomeric β-tubulin in these complexes. We compared the structures of β-tubulin in the native αβ-tubulin heterodimer and in the State I and State II complexes with TBCs (Fig. 2E). In the native tubulin heterodimer, the αβ-tubulin interface is stabilized by electrostatic and hydrophobic interactions (fig. S5E); the S9-H10 motif of β-tubulin, in particular, makes hydrophobic contacts with α-tubulin. Whereas the overall β-tubulin structure remains largely conserved [root mean square deviation (RMSD) ~1 Å], the S9-H10 motif undergoes a dramatic conformational rearrangement, with a maximal Cα displacement of ~18 Å owing to the loss of α-tubulin when β-tubulin forms a new interface with the AD3 region of TBCD (Fig. 2, E and F, and movie S2). The S9 loop (335 to 349), which is rich in bulky hydrophobic residues, moves near to the hydrophobic patch of AD3 (Fig. 2, E and F, and fig. S5G), thereby establishing a new interface that compensates for the absence of α-tubulin and minimizes the thermodynamic instability of the β-tubulin monomer. Incidentally, during its folding, the S9-H10 loop of β-tubulin is protected by the chaperonin TRiC/CCT (*10*), underscoring its inherent instability.

Collectively, the structures of the complexes in States I and II, along with supporting biochemical data, demonstrate the GTP-independent disassembly of αβ-tubulin and show that TBCD stabilizes the hydrophobic interface of β-tubulin after the heterodimer is disassembled.

GTP-induced conformational changes in TBC complex prime αβ-tubulin heterodimer assembly

GTP is a crucial regulator of tubulin folding and heterodimer assembly (*10*, *29*). Moreover, biochemical studies using hydrogen-deuterium exchange coupled with mass spectrometry have demonstrated that GTP exchange on Arl2 in the TBCD–Arl2–β-tub complex drives conformational changes in β-tubulin (*23*). To determine precisely how GTP functions in tubulin heterodimer assembly by TBCs, we analyzed the structures of States III and IV, containing TBCDE-Arl2 with β-tubulin monomer and with αβ-tubulin heterodimer, respectively, in the presence of GTP. Upon binding GTP, Arl2 undergoes a substantial reorientation of approximately 40 Å, transitioning from State II (apo form, compacted conformation) to State III (GTP-bound form, extended conformation) (Fig. 3, A and B, and fig. S6A). The N-terminal α-helix (H1, residues 1 to 14) of Arl2 remains anchored to TBCD, serving as a pivot during this transition (fig. S6A). The GTP-binding pocket undergoes a local conformational change, triggering reorganization of the salt bridge network within Arl2 (fig. S6B). This rearrangement of residues in the GTP-binding pocket of Arl2 facilitates a large-scale rearrangement in which the GTPase domain rotates ~180° and pivots ~106° around its fixed N-terminal α helix (Fig. 3B). These conformational changes alter Arl2's interaction surface with TBCD, shifting it from a β extension interaction between the β sheets of Arl2 and the N-terminal tail of TBCD to a hydrophobic interaction between Arl2 and the AD1 of TBCD [fig. S6C, (i)], whereas the principal interaction of the anchoring N-terminal α

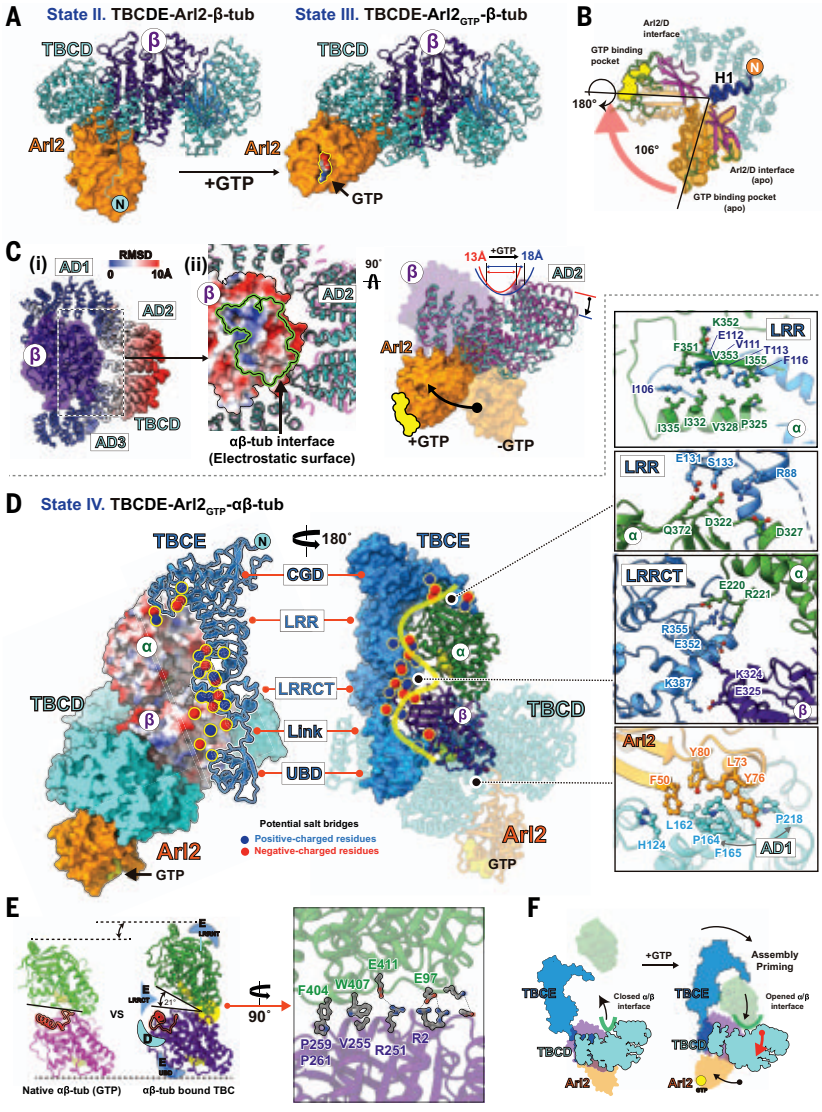


Fig. 3. GTP-induced conformational changes in TBCs prime αβ-tubulin heterodimer. (**A**) Cryo-EM structures illustrating Arl2 as a GTP-dependent switch. Upon GTP binding, Arl2 undergoes a major reorientation, pivoting around its N-terminal helix, which remains anchored to TBCD. (**B**) Summary diagram of Arl2 changes after GTP binding. The GTPase domain of Arl2 rotates ~180° relative to its fixed N-terminal helix and pivots by ~106°, reshaping its interface with TBCD. (**C**) Conformational rearrangement of AD2. (i) GTP binding induces a displacement of the AD2 domain of TBCD, (ii) increasing its distance from β-tubulin and relieving steric hindrance to allow α-tubulin access. AD1 and AD3 remain unchanged. (**D**) Architecture of the αβ-tubulin heterodimer bound to the TBCDE-Arl2 complex. Once GTP binds to Arl2 and AD2 is repositioned, α-tubulin binds the complex and TBCE engages both α- and β-tubulin through electrostatic complementarity and potential salt bridges across distinct surfaces. (**E**) Partially formed interface between α- and β-tubulin. Although the αβ-tubulin interface in this state is partially formed and not yet fully canonical, it is sufficiently stabilized by TBC interactions to maintain the complex. The opposite side is blocked by the LRR-CT of TBCE, which is opened by ~21° compared with the native interface (PDB 00007PJF). (**F**) Schematic of αβ-tubulin heterodimer assembly by the TBC complex. Without GTP, the heterodimer spontaneously dissociates upon TBC binding, leaving β-tubulin attached to TBCD. After GTP binding activates Arl2, AD2 opens, allowing α-tubulin to bind and prime the heterodimer for full assembly.

helix is unchanged. Thus, binding of GTP to Arl2 induces a profound conformational rearrangement in the protein.

The conformational rearrangement in Arl2 induced by binding GTP results in structural changes within TBCD. Specifically, the AD2 domain of TBCD is displaced by ~10 Å (Fig. 3C), which increases the distance between AD2 and β-tubulin from 13 to 18 Å. Although AD2 does not contact β-tubulin or Arl2 directly, its proximity to the αβ-tubulin interface suggests that it may regulate access of the α-tubulin for binding to β-tubulin. To test this hypothesis, we used molecular modeling to superimpose the native tubulin heterodimer structure onto the β-tubulin in the structures (fig. S6D). In State II, the position of AD2 partially occludes the α-tubulin in the heterodimer, thereby, potentially, hindering heterodimer assembly and stabilizing the complex in a disassembled state. By contrast, the increased spatial clearance observed in State III suggests that this steric hindrance is relieved, allowing access of α-tubulin to the interface. These findings suggest that AD2 regulates tubulin heterodimer assembly and disassembly, mediated by GTP-induced conformational transitions in Arl2.

In State IV, a major class of particles clearly resolves both α- and β-tubulin, and Arl2 has the same GTP-bound, extended conformation seen in State III (Fig. 3D). In this configuration, β-tubulin remains anchored to TBCD, whereas α-tubulin interacts primarily with TBCE. Unlike in State II, where TBCE has a flexible linker indicating a dynamic conformation (Fig. 2C), in State IV, TBCE has a rigid conformation indicating that its engagement with α-tubulin reduces its flexibility. TBCE's domain architecture creates two distinct binding grooves that accommodate both α- and β-tubulin. The LRR domain mainly associates with α-tubulin, whereas the UBD binds β-tubulin, and the LRR–C-terminal (CT) region binds both α- and β-tubulin by means of strong electrostatic complementarity.

Although the αβ-tubulin heterodimer in State IV contains multiple salt bridges and hydrophobic interactions, it does not have the fully canonical interface of the native heterodimer (Fig. 3E). The αβ-tubulin interface in State IV is notably more open than it is in the native tubulin heterodimer. The α-tubulin binding groove and the β-tubulin binding groove of TBCE are physically separated by the LRR-CT region, which is wedged between the monomers at the αβ-tubulin interface (Fig. 3, D and E), causing a 21° outward displacement pivoting around an electrostatic interaction in the αβ-tubulin interface (Fig. 3E). Whereas the α-tubulin structure closely resembles its native conformation, the S9-H10 motif of β-tubulin is extended and stabilized by TBCD's AD3 domain in a manner observed from State I (Fig. 2E). Consequently, we conclude that State IV likely represents a priming stage, where structural rearrangements of TBCs induced by GTP binding to Arl2 cooperatively position each tubulin monomer for correct heterodimer assembly (Fig. 3F).

Collectively, these findings demonstrate that GTP binding to Arl2 drives conformational changes in TBC complex that stabilize α- and β-tubulin, positioning them for correct heterodimer assembly. States III and IV, thus, are likely to be key intermediates in the GTP-dependent regulation of αβ-tubulin heterodimer assembly by Arl2 and TBCs, in which Arl2 serves as a molecular switch that steers the complex between the dissociation and association of the tubulin heterodimer.

TBCC completes assembly of the αβ-tubulin heterodimer

Previous biochemical studies indicated that TBCC facilitates tubulin heterodimer assembly cooperatively with GTP (*30*) and exclusively binds to the TBCD-Arl2 complex in the presence of GTP (*14*, *30*). Thus, we investigated how TBCC interacts with TBCDE-Arl2 and αβ-tubulin

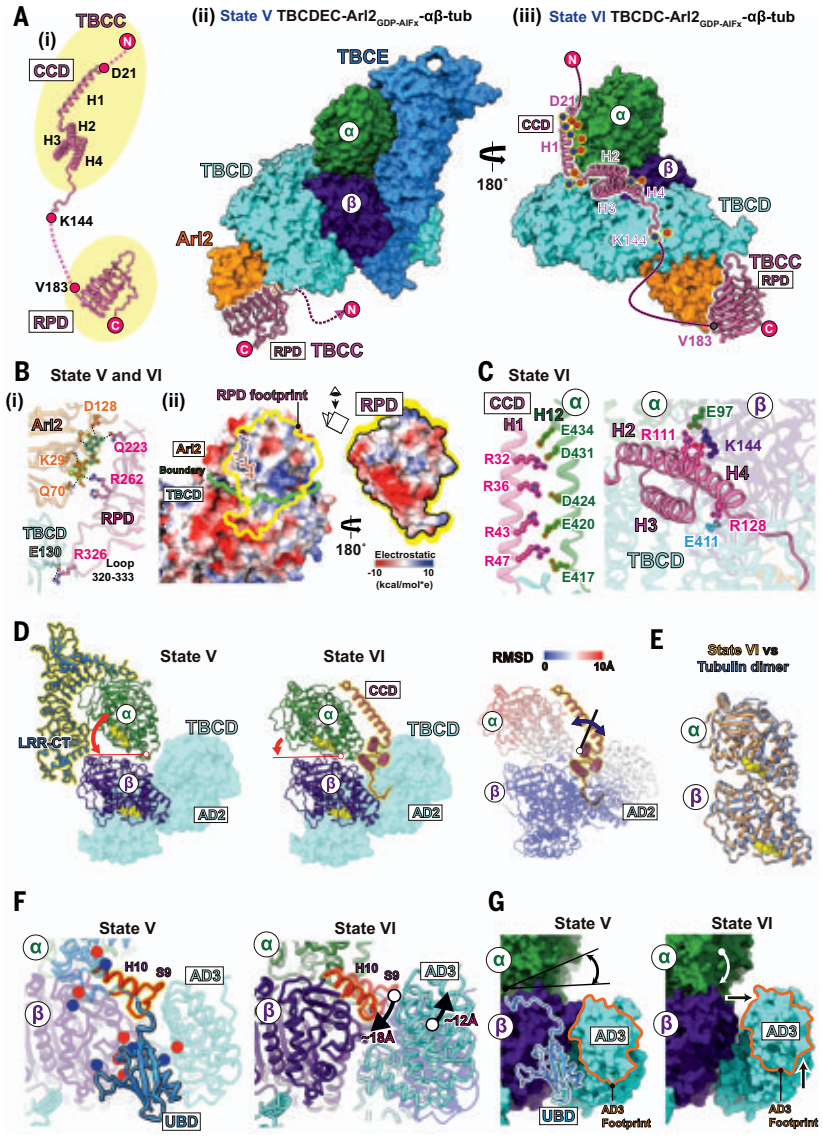


Fig. 4. TBCC completes assembly of the αβ-tubulin heterodimer. (**A**) Structural visualization of tubulin heterodimer assembly by the TBC complex and TBCC. (i) TBCC comprises a CCD and a β-sheet RPD, connected by a flexible linker. (ii) Cryo-EM structure of the TBCDEC-Arl2 complex bound to αβ-tubulin (State V). The TBCC RPD docks with Arl2 and TBCD. (iii) Cryo-EM structure of the TBCDC-Arl2 complex with αβ-tubulin (State VI), where the TBCC CCD clamps α- and β-tubulin, restoring the canonical tubulin interface. In State VI, TBCE detaches. (**B**) Binding interface between TBCC and Arl2. (i) A zoomed-in view of the Arl2 GTP-binding site, showing TBCC residues contributing to the catalytic core. (ii) Electrostatic surface view of TBCD-Arl2 illustrating the RPD binding interface, highlighting complementary surface charges and coverage around the GTP-binding pocket. The yellow footprint indicates TBCC-RPD, and the green line denotes the boundary between TBCD and Arl2, which together create the binding surface for RPD. (**C**) Interaction between the TBCC CCD and α-tubulin. A detailed view of the helix-to-helix contacts stabilizing α-tubulin within the complex. (**D**) Structural changes in the αβ-tubulin interface upon TBCC binding. In State V, α-tubulin is bound, forming an open interface with β-tubulin. Once the TBCC CCD clamps the heterodimer and TBCE detaches, the heterodimer adopts a native tubulin dimer interface with the reoriented β-tubulin S9-H10 motif. Structural alignment of αβ-tubulin in States V and VI reveals conformational remodeling at the dimer interface, associated with TBCC-CCD binding and TBCE release. (**E**) Superimposition of αβ-tubulin in State VI and the free tubulin heterodimer (PDB 00007PJF). Both heterodimers adopt the same orientation. (**F**) A zoomed-in view of β-tubulin motif remodeling. In State V, the β-tubulin S9-H10 motif flips outward, stabilized by AD3 of TBCD, whereas the TBCE UBD remains bound to the interface between TBCD AD3 and β-tubulin. In State VI, the motif is fully buried in the native dimer interface after TBCE detaches. The AD3 region of TBCD undergoes a reorientation, reducing curvature and releasing its interaction with the S9-H10 motif of β-tubulin as well, yielding a native heterodimer conformation. (**G**) Comparison of the UBD binding sites between State V and State VI. Disengagement of the β-tubulin S9-H10 motif from AD3 increases their distance, causing TBCE UBD detachment. Orange and white footprints represent the AD3 domain of TBCD and the UBD of TBCE, respectively. The white arrow indicates the downward movement of α-tubulin into the native dimer interface in State VI compared with State V, whereas the black arrows highlight the outward displacement of AD3 away from β-tubulin.

in the presence of GTP. Because TBCC induces rapid GTP hydrolysis by Arl2 with accompanying structural changes (*20*, *23*), instead of GTP, we used GTP with AlFx to induce a GDP-AlFx state, mimicking the transition state of GTP hydrolysis, thus minimizing structural heterogeneity and allowing high-resolution structural characterization. Upon adding TBCC and GTP with AlFx to TBCDE-Arl2 and αβ-tubulin, we uncovered two additional structural states: State V (TBCDEC–$Arl2_{GDP\text{-}AlFx}$–αβ-tub) and State VI (TBCDC–$Arl2_{GDP\text{-}AlFx}$–αβ-tub) (Fig. 4A). Of note, our maps resolved the posthydrolysis state, showing that GTP in Arl2 exhibited a cleaved density in the cryo-EM map corresponding to the AlFx replacing the γ-phosphate of GTP (figs. S2E, S3E, and S4E).

The overall architecture of State V is largely similar to that of State IV. The C-terminal RPD of TBCC is distinctly resolved near the GTP binding site of Arl2, whereas the N-terminal CCD is unresolved (Fig. 4A). Our structure reveals that GTP-induced conformational changes in Arl2 optimize the interface for the interaction with the TBCC-RPD, and a loop (residues 320 to 333) of the RPD directly interacts with residues in TBCD (Fig. 4B). Thus, GTP-bound Arl2 and TBCD together form the binding site for the RPD. This binding site is accessible only when Arl2 adopts its GTP-induced extended conformation, thus explaining the specific interaction of TBCC with Arl2 exclusively in the context of the TBCD-Arl2 complex in the presence of GTP. Also, our structure shows an arginine residue (R262) in the TBCC-RPD positioned adjacent to the phosphate analog of GDP-AlFx bound to Arl2 (Fig. 4B), This residue, R262, is homologous to R186 in yeast TBCC, a conserved residue essential for GTPase activating protein (GAP) activity, as mutations abolish its function (*14*). These findings suggest that TBCC-RPD contributes directly to the catalytic environment necessary for efficient Arl2 GTP hydrolysis.

In State VI, both the CCD and RPD of TBCC are clearly resolved (Fig. 4A). The RPD bound to Arl2 has the extended conformation observed in State V, whereas the CCD interacts primarily with α-tubulin and also contacts β-tubulin and TBCD. The H1 helix of the CCD forms multiple salt bridge interactions with the H12 helix of α-tubulin, and the H2 to H4 helices of the CCD bundle together to span the interface between α- and β-tubulin and TBCD (Fig. 4C). No density corresponding to TBCE was detected in State VI, suggesting that binding of the CCD of TBCC and binding of TBCE to α-tubulin are mutually exclusive. The CCD effectively acts as a molecular clamp to stabilize the interface among the TBCD and α- and β-tubulin complex while facilitating the release of TBCE from the complex. The αβ-tubulin dimer interface in State VI is

identical to the canonical native tubulin heterodimer interface (Fig. 4, D and E). In State VI, the S9-H10 motif of β-tubulin folds back into the αβ-tubulin dimer interface, restoring the native hydrophobic contacts and completing dimer maturation (Fig. 4F). This conformational transition buries the LRR-CT binding site and triggers TBCE release. Concurrently, the motif disengages from TBCD's AD3 domain, which undergoes reorientation and curvature relaxation. These changes relieve steric hindrance and expose the surface area previously occupied by the UBD of TBCE (Fig. 4G).

TBCD functions as a GAP for β-tubulin

The free αβ-tubulin heterodimer has a low GTPase activity because it lacks the catalytic contribution from neighboring tubulin heterodimers in the polymeric microtubule (*31*). Yet newly assembled αβ-tubulin heterodimers are released from TBC complexes upon hydrolysis of GTP by β-tubulin (*11*). This suggests that TBCs may function as GAPs for β-tubulin. To investigate which component(s) of the TBC complex might provide this GAP activity, we compared our structures with the interface between heterodimers within the microtubule lattice (PDB 00006DPU) (*32*). The α-tubulin of one heterodimer functions as a GAP for the β-tubulin in the adjacent heterodimer (Fig. 5A) (*33*). Specifically, the two negatively charged residues D251 and E254 in α-tubulin are suggested to contribute to the catalytic core required for GTP hydrolysis by β-tubulin by coordinating water molecules (*7*). We looked for a similar arrangement of residues in the TBC-tubulin complex. Indeed, TBCD encircles β-tubulin, covering the GTP-binding pocket and positioning its residues near the γ-phosphate of GTP bound in β-tubulin (Fig. 5B). In our structures, we found the charged residues D65 and E69 in TBCD at its interface with β-tubulin based on the location of resolved densities in our cryo-EM maps attributable to an Mg^{2+} ion and water molecules (Fig. 5B). We hypothesize that these TBCD residues may mimic the roles of D251 and E254 in α-tubulin by coordinating water molecules, thereby facilitating GTP hydrolysis by β-tubulin.

To test whether TBCD D65 and E69 contribute to β-tubulin's GTP hydrolysis, we generated a double alanine mutant of TBCD (Fig. 5C). To distinguish GTP hydrolysis by β-tubulin from that by Arl2, we introduced an Arl2 Q70L mutation that prevents its GTP hydrolysis while GTP remains bound (*29*). We then purified the following four TBCDE-Arl2 complexes—wild-type (WT) TBCDE-Arl2, TBCD(D65A/E69A)E-Arl2, TBCDE-Arl2(Q70L), and TBCD(D65A/E69A)E-Arl2(Q70L)—and measured their GTPase activity using a malachite green assay in the presence of αβ-tubulin (Fig. 5, C and D, and fig. S7). Consistent with prior observations, TBCDE-Arl2 and αβ-tubulin, in the absence of TBCC, exhibited very low GTPase activity (Fig. 5D) (*14*). In the presence of TBCC, WT TBCDE-Arl2 had the highest GTPase activity owing to the combined activity of Arl2 and β-tubulin (Fig. 5D). TBCD(D65A/E69A)E-Arl2 had ~25% less GTPase activity than the WT complex, indicating that residues D65 and E69 of TBCD play a direct role in GTPase hydrolysis by β-tubulin, although Arl2 contributes a substantial part of the overall GTPase activity of the complex (Fig. 5D). Unexpectedly, TBCDE-Arl2(Q70L) also had little GTPase activity, similar to that of TBCD(D65A/E69A)E-Arl2(Q70L) (Fig. 5D). This result suggests that the Arl2(Q70L) mutation directly impacts β-tubulin's GTPase activity, indicating a functional dependence of GTP hydrolysis by β-tubulin on the catalytic activity of Arl2.

Discussion

By solving the cryo-EM structures of six complexes containing human TBCs and monomeric β-tubulin or heterodimeric αβ-tubulin in the presence and absence of GTP, we elucidated the mechanisms by which TBCs guide the association of α- and β-tubulin monomers into native αβ-tubulin heterodimers and, conversely, the dissociation of heterodimers into monomers (Fig. 6, A and B). The six complexes reflect distinct states of tubulin heterodimer assembly and disassembly. The State I (TBCD–Arl2–β-tub) and State II (TBCDE–Arl2–β-tub) structures, in which the αβ-tubulin heterodimer has been disassembled, show β-tubulin bound to TBCs in the absence of GTP. They demonstrate that disassembly of αβ-tubulin is independent of GTP and show that TBCD shields the exposed hydrophobic surfaces of β-tubulin after the heterodimer is disassembled to prevent misfolding or aberrant interactions.

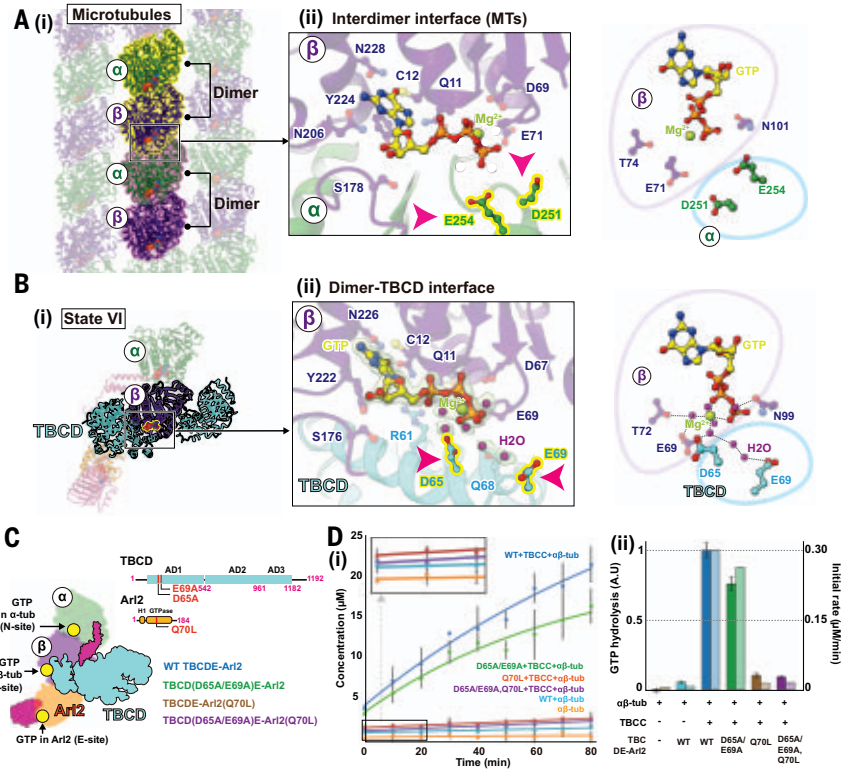


Fig. 5. TBCD functions as a GAP for β-tubulin. (**A**) α-Tubulin functions as a GAP for β-tubulin in the microtubule (MT) lattice (PDB 00006DPU). (i) Interdimer interface within the microtubule, highlighting the β-tubulin GTP site adjacent to α-tubulin. (ii) Detailed view of β-tubulin's GTP binding pocket and neighboring α-tubulin residues forming the catalytic core. Highlighted residues, αD251 and αE254, indicate key contributors to GTP hydrolysis at the interdimer interface. (**B**) TBCD recapitulates α-tubulin's GAP function for β-tubulin within the TBC complex. (i) The TBCDC-Arl2 complex with an assembled αβ-tubulin heterodimer in a native interface. (ii) Close-up view of the β-tubulin GTP binding site, highlighting the coordinated ion and water molecules, along with the two negatively charged TBCD residues that contribute to the catalytic environment. Residues and coordinated waters from β-tubulin and TBCD involved in GTP hydrolysis. (**C**) Schematic depicting GTP binding sites in the TBC–αβ-tub complex. Constructs for Arl2 and TBCD highlight the catalytic residues and mutant sites used in assays. (**D**) GTP hydrolysis measured through malachite green assays. (i) Phosphate concentrations in the solution were quantified and plotted as a time series. The black-outlined box shoes a zoomed-in of the lower portion at early time points. (ii) Absorbance at 621 nm after 80 min was normalized and plotted (left bars). Initial reaction rates (micromolar per minute), derived from model fitting, for each dataset are shown (right bars). Error bars represent mean ± 95% confidence intervals ($n = 3$ independent experiments). A.U, arbitrary units.

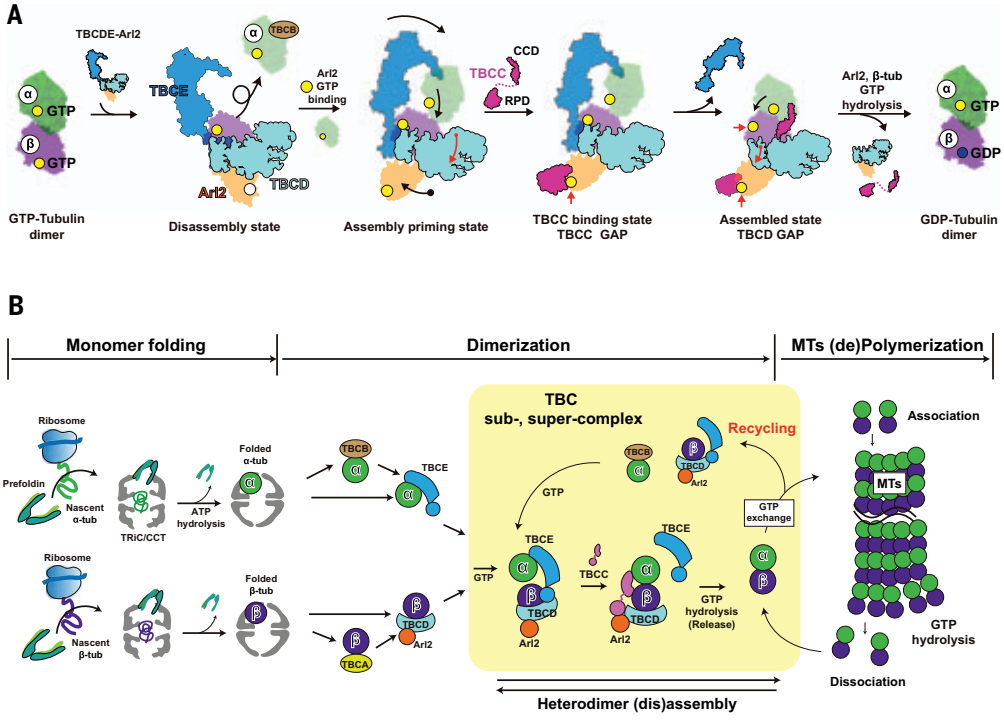


Fig. 6. Roles of TBCs in tubulin folding, assembly, recycling, and quality control. (**A**) Schematic of αβ-tubulin heterodimer dissociation and assembly. Upon binding the TBC complex, the heterodimer dissociates, leaving β-tubulin bound while α-tubulin is released (often sequestered by TBCB). GTP binding on Arl2 triggers conformational changes in TBCD, allowing TBCE to capture α-tubulin and prime αβ-tubulin reassembly. TBCC binding promotes TBCE release and completes αβ-tubulin heterodimer formation. Lastly, β-tubulin hydrolyzes GTP, aided by TBCD's GAP activity, releasing a mature GDP–αβ-tubulin heterodimer. (**B**) A comprehensive framework for tubulin biogenesis includes folding, assembly, recycling, and quality control. TBCs help release newly folded monomers from TRiC/CCT, delivering them to TBCD and TBCE. TBCC then promotes final maturation by activating Arl2's GTPase, enabling the release of fully assembled αβ-tubulin heterodimers. The TBCDE-Arl2 complex underlies both de novo assembly and heterodimer recycling.

Moreover, Arl2 switches the TBC complex between dissociation and association of the heterodimer. In State III (TBCDE–Arl2$_{GTP}$–β-tub) and State IV (TBCDE–Arl2$_{GTP}$–αβ-tub), GTP binding to Arl2 drives conformational changes in TBCD and TBCE that guide α- and β-tubulin toward a partially assembled heterodimer. These two states are likely key intermediates in the GTP-dependent regulation of tubulin assembly by Arl2. In the presence of TBCC in State V (TBCDEC–Arl2$_{GDP\text{-}AlFx}$–αβ-tub) and State VI (TBCDC–Arl2$_{GDP\text{-}AlFx}$–αβ-tub), in which we used GTP with AlFx to mimic the posthydrolysis state of GTP, the C-terminal RPD of TBCC binds near the GTP binding site of Arl2 and facilitates GTP hydrolysis by Arl2, whereas the N-terminal CCD interacts primarily with α-tubulin and acts as a molecular clamp to stabilize the interface among the TBCD and α- and β-tubulin complex, allowing formation of the mature, native heterodimer structure and facilitating the release of TBCE from the complex. Lastly, we used mutagenesis to show that TBCD contributes key catalytic residues (D65 and E69) that coordinate the water molecules essential for GTP hydrolysis by β-tubulin in the TBC complex, allowing the release of mature GDP–αβ-tubulin heterodimers from the complex.

This detailed description of the molecular basis of TBC function at atomic resolution underpins a comprehensive understanding of the tubulin heterodimer lifecycle from de novo folding of α- and β-tubulin monomers through assembly of new heterodimers and recycling of existing heterodimers. TBCA and TBCB facilitate release from tubulin monomers from TRiC/CCT (*17*) and deliver them to the downstream cofactors TBCD-Arl2 and TBCE (*11*, *22*, *34*). TBCC then promotes maturation of the heterodimers by activating the GTPase activity of Arl2 (*14*), leading to the release of fully assembled αβ-tubulin heterodimers. Our structural data reveal that the TBCDE–Arl2 complex lies at the heart of a unified mechanism that governs both de novo tubulin assembly and heterodimer recycling. Recent in situ cryo–electron tomography studies have identified TBC-like particles within the lumen of microtubules, suggesting that TBCs may mediate tubulin disassembly and reassembly locally in regions of high curvature or lattice defects (*35*). By serving as a local reservoir for tubulin monomers and facilitating removal of damaged subunits, TBCs might thus contribute to the spatial regulation of microtubule homeostasis, particularly in peripheral regions of the cell. Taken together, these findings broaden the conventional view of TBC function, establishing them as key regulators of both de novo tubulin assembly and heterodimer recycling, two processes that converge on a common set of structural transitions in TBCDE-Arl2.

GTP hydrolysis is fundamental to microtubule dynamics, driving the transition between GTP- and GDP-tubulin, which governs microtubule polymerization, depolymerization, and dynamic instability, key processes essential for intracellular transport, mitosis, and cytoskeletal remodeling (*36*). GTP binding to tubulin likely occurs during the folding of α- and β-tubulin monomers by TRiC/CCT, stabilizing folding intermediates and facilitating proper tubulin maturation (*10*). Beyond its role in folding and microtubule dynamics, however, our study reveals a further role for GTP in tubulin heterodimer assembly and disassembly, with the GTP-binding protein Arl2 the master regulator in the TBC complex. Arl2 likely plays a dedicated and nonredundant role in tubulin heterodimer biogenesis, as its knockout disrupts microtubule dynamics (*14*, *37*, *38*), and its specific interaction surface with TBCD (Figs. 1D and 3, A and B, and fig. S3, A and B) is not conserved among other ADP ribosylation factor family members. Arl2 cycles between apo- and GTP-bound states to switch the function of the TBC complex between assembly and disassembly of the αβ-tubulin heterodimer. In its GTP-bound state, Arl2 directs TBC-mediated assembly, whereas, in the absence of GTP, it stabilizes the disassembled state, preventing premature heterodimer formation. As cellular GTP concentrations are relatively stable under physiological conditions (*39*), this Arl2-mediated regulatory cycle likely operates constitutively as a housekeeping mechanism, continuously supplying tubulin heterodimers competent to support microtubule growth, repair, and turnover. Thus, GTP is indispensable throughout the tubulin life cycle, regulating monomer folding, heterodimer assembly, recycling, and microtubule function to ensure proper cytoskeletal maintenance and cellular homeostasis.

Whereas the RPD of TBCC functions as a GAP for Arl2 (*14*), we demonstrate also that TBCD acts as a GAP for β-tubulin, providing negatively charged catalytic residues that position a water molecule near the β-tubulin–bound GTP for hydrolysis. GTP hydrolysis by β-tubulin

is crucial for release of the newly assembled heterodimer (*11*), and we show that it is functionally coupled to Arl2's GTPase activity, establishing a two-tier, GTP-driven mechanism that governs the maturation and quality-controlled release of αβ-tubulin heterodimers. In this mechanism, TBCC facilitates dimer maturation and, through its GAP activity on Arl2, enables the subsequent GTP hydrolysis of β-tubulin by TBCD, which serves as a key quality control checkpoint to ensure the release of functionally competent heterodimers.

The TBC-mediated cyclical assembly and disassembly of tubulin heterodimers provides a stringent quality control mechanism, preventing defective tubulin from disrupting microtubule function. We note that β-tubulin remains GTP-bound in our structures, suggesting that an additional, not-yet-observed structural transition may be necessary to activate its GTPase function in β-tubulin. Future studies will be required to determine the precise trigger for this transition and whether specific TBCs or conformational rearrangements facilitate β-tubulin's GTP hydrolysis.

Materials and methods

Plasmid construction

Human TBCA, Human TBCB: Human *TBCA* (UniProt ID: O75347) and *TBCB* (UniProt ID: Q99426) genes were amplified from plasmids purchased from DNASU plasmid repository. *TBCA* was amplified using the 5′ primer cgccatatggccgatcctcgcg and the 3′ primer ggattcagtgaagttagaagcctgactcgagcgg, while *TBCB* was amplified using the 5′ primer cgccatatggaggtgacggggg and the 3′ primer actacgggttggacgagatatgactcgaggcg. The amplified genes were cloned into the pET-29b(+) expression vector using conventional cloning methods. The resulting constructs, pET-29b(+)-TBCA and pET-29b(+)-TBCB, each contains a C-terminal hexa-histidine tag.

Human TBCC: Human *TBCC* (UniProt ID: Q15814) was amplified from plasmids purchased from Twist Bioscience using the 5′ primer ccggaccggtgccaccatgcaccatcaccaccatca cggttcaggagagtccgtcagt and the 3′ primer ggccgcggccgcttagtcccactggatatttcgctcctcttcagg aagaatactccagtttgg. The amplified product was cloned into the EEV-CAG-MCS-WPRE-pA mammalian expression vector using AgeI and NotI restriction sites. The resulting construct, EEV-CAG-MCS-WPRE-pA-TBCC, contains an N-terminal hexa-histidine tag.

Human TBCD (WT), Human TBCE: Human *TBCD* (UniProt ID: Q9BTW9) and *TBCE* (UniProt ID: Q15813) were amplified from plasmids purchased from Twist Bioscience. *TBCD* was amplified using the 5′ primer ccggaccggtgccaccatgtggagcca and the 3′ primer tgcctgctgagcggccgcggcc, while *TBCE* was amplified using the 5′ primer ccggaccggtgccaccatgtggagcca and the 3′ primer tattagtgcgatggtgagcggccgcggcc. The amplified products were cloned into the EEV-CAG-MCS-WPRE-pA mammalian expression vector using AgeI and NotI restriction sites. The resulting constructs, EEV-CAG-MCS-WPRE-pA-TBCD and EEV-CAG-MCS-WPRE-pA-TBCE, each contains an N-terminal StrepII-tag.

Human TBCD (D65A, E69A): The human *TBCD* (D65A, E69A) mutant was generated by PCR-based mutagenesis with human *TBCD* (WT) as the template. Mutagenesis was performed with the 5′ primer gtaataatggccaaataccaggcgcagcctcatctgttggacccg and the 3′ primer atgaggctgcgcctggtatttggccattattacgcggaaccgctc.

Human Arl2 (WT and Q70L): Human *Arl2* (WT, UniProt ID: P36404) was amplified from plasmids purchased from Twist Bioscience using the 5′ primer ccggaccggtatggggctcctgaccatt and the 3′ primer ggccgcggccgctcagtcagctgt. The amplified product was cloned into the EEV-CAG-MCS-WPRE-pA vector using AgeI and NotI. The *Arl2 (Q70L)* mutant was generated through PCR-based mutagenesis using *Arl2 (WT)* as the template, using the 5′ primer gtgggtggcctgaagtccctgcggtcctactggcggand the 3′ primer cagggacttcaggccacccacatcccagatgttc agc.

Protein expression and purification

Human TBCA, Human TBCB: Recombinant plasmids pET-29b(+)-TBCA and pET-29b(+)-TBCB were individually transformed into the *E. coli* BL21(DE3) expression host (Enzynomics, CP111). The cells were cultured in 1 L of LB medium supplemented with 100 μg/mL kanamycin at 37°C until the OD_{600} reached 0.6. Protein expression was induced by adding 1 mM isopropyl-β-D-thiogalactoside (IPTG) to the medium. After overnight incubation at 18°C, the cells were harvested by centrifugation and resuspended in 20 mL of lysis buffer (50 mM HEPES pH 7.0, 100 mM KCl, 1 mM $MgCl_2$ and 1 mM Dithiothreitol (DTT)), supplemented with a protease inhibitor cocktail tablet (Roche). The resuspended cells were lysed by sonication, and cell debris was removed by centrifugation at 20,000 × g for 30 min at 4°C. His-tagged TBCA and TBCB proteins were purified using 2.5 mL Ni-NTA agarose resin (Thermo Fisher Scientific). After protein binding, the resin was washed with 15 mL of wash buffer containing 50 mM HEPES pH 7.0, 100 mM KCl, 1 mM $MgCl_2$, 20 mM imidazole pH 8.0, and 1 mM DTT. Proteins were eluted with 15 mL of elution buffer containing 50 mM HEPES pH 7.0, 100 mM KCl, 1 mM $MgCl_2$, 500 mM imidazole pH 8.0, and 1 mM DTT. The eluted proteins were concentrated to 500 μL using Amicon Ultra Centrifugal filter (10 kDa MWCO) and further purified using a Superose-6 size exclusion column equilibrated with buffer containing 50 mM HEPES pH 7.0, 100 mM KCl, 1 mM $MgCl_2$ and 1 mM DTT. TBCA- and TBCB-containing fractions were identified by SDS-PAGE, concentrated, aliquoted, snap frozen, and stored at -80°C.

Human TBCC: Human TBCC plasmid was transfected into HEK293S GnTI- (ATCC, CRL-3022) cells. Cells were incubated at 37°C for ~72 hours, harvested by centrifugation at 4000 rpm, and then lysed in lysis buffer (50 mM MES pH 6.0, 100 mM KCl, 3 mM β -mercaptoethanol) supplemented with a protease inhibitor cocktail tablet (Roche). The lysates were cleared by centrifugation at 15,000rpm for 1 hour at 4°C. The resulting supernatant was filtered through a 0.8 μm minisart syringe filter and then subjected to affinity-purification using a gravity column with TALON Metal Affinity Resin (Takara). The resin was washed with wash buffer (50 mM MES pH 6.0, 100 mM KCl, 3 mM β-mercaptoethanol) and proteins were eluted with a three-column volume of wash buffer containing 400 mM imidazole. The eluted protein-containing fractions were concentrated using Amicon Ultra Centrifugal filter (30 kDa MWCO) and further purified by size-exclusion chromatography on a superdex200 increase 10/300 GL column equilibrated with wash buffer. Fractions containing TBC proteins were individually identified by SDS-PAGE, concentrated, aliquoted, snap frozen, and stored at -80°C.

Human TBC-DE-Arl2 and TBC-DE-Arl2 mutants: TBCD, TBCE, and Arl2 plasmids were transfected into HEK293S GnTI- (ATCC, CRL-3022) cells at a ratio of 1:1:1. Cells were incubated at 37°C for ~72 hours, harvested by centrifugation at 4000 rpm, and then lysed in lysis buffer (150 mM KCl, 50 mM HEPES pH 7.0, 1 mM $MgCl_2$, 50 μM GTP, 3 mM β-mercaptoethanol) supplemented with a protease inhibitor cocktail tablet (Roche). The lysates were cleared by centrifugation at 15,000rpm for 1 hour at 4°C. The resulting supernatant was filtered through a 0.8 μm minisart syringe filter and then subjected to affinity-purification using a gravity column with Strep-Tactin Sepharose resin (IBA). The resin was washed with low-salt buffer (100 mM KCl, 50 mM HEPES pH 7.0, 1 mM $MgCl_2$) and proteins were eluted with a five-column volume of low-salt buffer containing 2.5 mM desthiobiotin. The eluted TBC-DE-Arl2 containing fractions were concentrated using Amicon Ultra Centrifugal filter (50 kDa MWCO) and further purified by size-exclusion chromatography on a Superose6 column equilibrated with low-salt buffer. Fractions containing TBC-DE-Arl2 protein were identified by SDS-PAGE, concentrated, aliquoted, snap frozen, and stored at -80°C.

Tubulin: Tubulin protein powder (>99% pure) purified from porcine brain (Cytoskeleton Inc., T240-D) was passed over a Superose-6 size

exclusion column equilibrated with buffer containing 50 mM HEPES pH 7.0, 100 mM KCl, 1 mM $MgCl_2$ and 1 mM DTT. Tubulin containing fractions were pooled, concentrated, aliquoted, snap frozen, and stored at -80°C for long-term storage.

Native PAGE and western blotting analysis

TBCs were mixed with 10 μM tubulin dimers at adjusted volumes to ensure stoichiometric ratios. The reactions were then incubated at 30°C for 30 min, with or without 1 mM GTP. The samples were loaded onto a non-denaturing polyacrylamide gel containing 0.1 M MES-NaOH pH 6.7, 1 mM $MgCl_2$, 1 mM EGTA, 1 mM GTP, 0.14% APS, 0.01% TEMED, and 6% acrylamide. Electrophoresis was carried out at 4°C in a buffer containing 0.1 M MES-NaOH pH 6.7, 1 mM $MgCl_2$, and 1 mM EGTA. Proteins were transferred onto nitrocellulose membranes and blocked with a 2.5% skim milk solution in TBST buffer (20 mM Tris-HCl, pH 7.4; 150 mM NaCl; 0.5% (v/v) Tween 20) for 1 hour at room temperature. Membranes were then incubated overnight with primary antibodies: anti-α tubulin (1:5000, Thermo Fisher Scientific, 62204), anti-β tubulin (1:5000, Proteintech, 66240-1-Ig), anti-His (1:5000, Thermo Fisher Scientific, MA1-21315), and anti-Strep II (1:5000, Abcam, ab184224). After washing, the membranes were incubated with a secondary antibody (Donkey anti-mouse, 1:10,000, Abcam, AB175774) for 1 hour. Protein bands were visualized using a ChemiDoc imaging system (Amersham ImageQuant 800, Cytiva).

Mass spectrometry analysis

A Thermo Fisher Scientific Quadrupole-Orbitrap instrument (Thermo Fisher Scientific) equipped with Dionex U 3000 RSLCnano HPLC system was used. Mass spectrometric analyses were performed using a Thermo Fisher Scientific Orbitrap Exploris 240 mass spectrometer. Fractions were reconstituted in solvent A (Water/Acetonitrile (98:2 v/v), 0.1% Formic acid) and then injected into LC-nano ESI-MS/MS system. Samples were first trapped on a Acclaim PepMap 100 trap column (100μm x2cm, nanoViper C18, 5μm, 100 Å, Thermo Fisher Scientific, part number 164564) and washed for 6 min with solvent A (water/ACN (98:2 v/v), 0.1% Formic acid at a flow rate of 4 μL/min, and then separated on a PepMap RSLC C18 column (75μm × 15cm, nanoViper C18, 3 μm, 100 Å, Thermo Fisher Scientific, part number ES900) at a flow rate of 300 nL/min. The LC gradient was run at 2% to 8% solvent B over 10min, then from 8% to 30% over 55min, followed by 90% solvent B (100% ACN and 0.1% Formic acid) for 4min, and finally 2% solvent B for 20min. Xcalibur software version 4.4 was used to collect MS data. The Orbitrap analyzer scanned precursor ions with a mass range of 350-1,800 m/z with 60,000 resolutions at m/z 200. Mass data are acquired automatically using proteome discoverer 2.5 (Thermo Fisher Scientific)

Size-exclusion chromatography

TBCDE-Arl2 complex and αβ-tubulin heterodimer were mixed at 1:1.1 stoichiometry and incubated for 30 min at room temperature. Both protein complexes were purified as homogeneous complexes as described above. The mixture was applied onto a Superdex 200 increase 10/300 GL column (Cytiva) equilibrated with low-salt buffer (100 mM KCl, 50 mM HEPES pH 7.0, 1 mM $MgCl_2$). Eluted fractions were analyzed by SDS-PAGE and Western blotting as described above.

GTP hydrolysis assay

To compare the GTP hydrolysis activity among the TBCDE-Arl2 mutants, assays were performed using (i) wild-type (WT) TBCDE-Arl2, (ii) TBCD(D65A, E69A)E-Arl2, (iii) TBCDE-Arl2(Q70L), and (iv) TBCD(D65A,E69A)E-Arl2(Q70L). GTP hydrolysis activity was measured using the Malachite Green Assay Kit (Sigma-Aldrich, MAK307) following the manufacturer's protocol. Reactions were prepared by mixing proteins in low-salt buffer (100 mM KCl, 50 mM HEPES pH 7.0, 1 mM $MgCl_2$) to achieve concentrations of 0.4 μM TBCDE-Arl2, 0.8 μM TBCC, 0.8 μM tubulin, and 0.8 μM TBCA. Reactions were plated in a 96-well plate (3595; Corning) at 40 μL per well, along with a 0–40 μM phosphate standard. 0.1 mM GTP (Sigma-Aldrich, 10106399001) was added at 10-min intervals to a final concentration of 50 μM. Samples were then incubated at 30°C to monitor GTP hydrolysis over time. Following incubation, Malachite Green reagent was added, and the plate was incubated at room temperature for 30 min. Absorbance at 621 nm was measured using a plate reader (Tecan i-control). Inorganic phosphate concentrations were determined using a 0–40 μM linear phosphate standard curve prepared under the same reaction conditions. The experimental data were fitted to a one-phase association curve.

Cryo-EM grid preparation and data acquisition

TBCD–Arl2–β-tub and TBCDE–Arl2–β-tub (States I and II): The TBCDE-Arl2 complex at a concentration of 1 mg/mL was incubated with αβ-tubulin dimers, TBCA and TBCB at a 1:1.2:1.2:1.2 molar ratio for 30 min at 30°C. A 3 μL aliquot of the complex was applied to glow-discharged 300-mesh UltrAuFoil R1.2/1.3 holey gold grids. Grids were blotted for 2.5-3 s under 100% humidity at 20°C and then plunge-frozen into liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific) at the Center for Macromolecular and Cell Imaging (CMCI), Seoul National University (SNU). A total of 15,623 movies were collected in counted mode using a Titan Krios (Thermo Fisher Scientific, IMP at POSTECH) equipped with a Gatan K3 direct electron detector and a BioQuantum energy filter with slit width 18 eV. All the movies were acquired with a defocus range of -0.8 to -2.0 μm, at a pixel size of 0.624 Å, and a total electron dose of 50 $e^-/Å^2$.

TBCDE–$Arl2_{GTP}$–β-tub and TBCDE–$Arl2_{GTP}$–αβ-tub (States III and IV): The TBCDE-Arl2 complex at a concentration of 1.2 mg/mL was incubated with αβ-tubulin dimers and TBCA at a 1:2:2 molar ratio with 1 mM GTP for 30 min at 30°C. A 3 μL aliquot of the complex was applied to glow-discharged 300-mesh UltrAuFoil R1.2/1.3 holey gold grids. Grids were blotted for 2.5 s under 100% humidity at 20°C and then plunge-frozen into liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific) at the Center for Macromolecular and Cell Imaging (CMCI), Seoul National University (SNU). A total of 17,543 movies were collected in electron-event representation (EER) mode using a Titan Krios (Thermo Fisher Scientific, NCIRF at SNU) equipped with a Falcon 4i direct electron detector and a Selectris X energy filter (Thermo Fisher Scientific) with slit width 10 eV. All the movies were acquired with a defocus range of -0.8 to -2.0 μm, at a pixel size of 0.74 Å, and a total electron dose of 50 $e^-/Å^2$.

TBCDEC–$Arl2_{GDP\text{-}AlFx}$–αβ-tub and TBCDC–$Arl2_{GDP\text{-}AlFx}$–αβ-tub (States V and VI): The TBCDE-Arl2 complex at a concentration of 1.6 mg/mL was incubated with αβ-tubulin, TBCA, and TBCC at a molar ratio of 1:1:1.2:3 under 1 mM GTP, 1 mM $Al(NO_3)_3$, and 6 mM NaF conditions for 30 min at 30°C. A 3 μL aliquot of the complex was applied to glow-discharged 300-mesh UltrAuFoil R1.2/1.3 holey gold grids. Grids were blotted for 2.5 s under 100% humidity at 20°C and then plunge-frozen into liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific) at the Center for Macromolecular and Cell Imaging (CMCI), Seoul National University (SNU). A total of 43,925 movies were collected in two separate datasets in electron-event representation (EER) mode using a Titan Krios (Thermo Fisher Scientific, NCIRF at SNU) equipped with a Falcon 4i direct electron detector and a Selectris X energy filter (Thermo Fisher Scientific) with slit width 10 eV. All the movies were acquired with a defocus range of -0.8 to -2.0 μm, at a pixel size of 0.74 Å, and a total electron dose of 50 $e^-/Å^2$.

Data processing and 3D refinement

All image processing was done in CryoSPARC v4.5, v4.6 (*40*, *41*). Computing resources were utilized at CMCI at Seoul National University. The detailed processing parameters are summarized in table S1.

TBCD–Arl2–β-tub and TBCDE–Arl2–β-tub (States I and II): Movies were aligned using patch motion correction, and CTF parameters were estimated through patch CTF estimation in CryoSPARC. Initial particles were picked and extracted using template-based particle picking in CryoSPARC, with preliminary 2D classes as reference. After multiple rounds of 2D classifications and 2D class selections, an initial set of 506,791 particles were refined using non-uniform refinement. This refined particle set was used as a reference for Topaz training, and 2,347,984 particles were newly extracted from the original micrographs through Topaz extraction. After heterogeneous refinement, 30.8% of extracted particles were selected and a non-uniform refinement was performed to these particles. Using the refined volume as a reference, the particles underwent reference-based motion correction. CTF refinement and non-uniform refinement were performed on 714,398 particles, generating a consensus map of State II at 2.14 Å resolution. The TBCE UBD was masked, and 3D classification was performed on these particles. Particles were classified into six classes, including class 0, which lacked TBCE UBD density. Using these particles, a consensus map at 2.48 Å resolution was generated for State I through non-uniform refinement. For TBCE LRR domain density, heterogeneous refinement and 3D classification with a TBCE mask were performed on the subset of 714,398 particles above. Further local refinements and additional 3D classification refined TBCE LRR domain density, represented by 68,994 particles, at resolution of 3.59 Å. Consensus maps of State I, II and focused map of TBCE LRR domain were sharpened using DeepEMhancer (*42*). The composite map for State II was obtained by combining the focused map of the TBCE LRR domain with the State II consensus map using vop maximum command in UCSF ChimeraX (*43*).

TBCDE–Arl2$_{GTP}$–β-tub and TBCDE–Arl2$_{GTP}$–αβ-tub (States III and IV): Movies were aligned using patch motion correction, and CTF parameters were estimated through patch CTF estimation in CryoSPARC. Initial particles were picked and extracted using template-based particle picking in CryoSPARC, with preliminary 2D classes as reference. After multiple rounds of 2D classifications followed by a single heterogeneous refinement, an initial set of 155,246 particles was refined using non-uniform refinement. This refined particle set was used as a reference for Topaz training, and 1,314,052 particles were newly extracted from the original 17,543 micrographs through Topaz extraction. Heterogeneous refinement classified the topaz extracted particles into six subsets. From these subsets, class 2 (18.0%) was selected and subjected to non-uniform refinement, generating State IV consensus map of 237,104 particles at 2.55 Å resolution. The remaining classes underwent heterogeneous refinement, and 515,321 particles were selected for two cycles of non-uniform refinement, generating State III consensus map at 2.45 Å resolution. For focused refinement of State III, one mask was created for Arl2, while another encompassed the TBCD AD3 domain and TBCE UBD. Local refinements were performed on these masked regions, generating 2.56 Å and 2.41 Å resolution local-refined maps for each domain. Similarly, for State IV, one mask was created for Arl2, while another covered α-tubulin, the TBCE LRR domain, and TBCE UBD. Local refinements were performed on these masked regions, generating 2.70 Å resolution local-refined maps for each domain. The final maps for State III and State IV were obtained by combining the focused maps with the consensus maps using Phenix combine-focused-maps and sharpened with DeepEMhancer.

TBCDEC–Arl2$_{GDP\text{-}AlFx}$–αβ-tub and TBCDC–Arl2$_{GDP\text{-}AlFx}$–αβ-tub (States V and VI): Two independent datasets were collected, comprising 26,616 and 17,309 movies, respectively. For each dataset, movies were aligned using patch motion correction, and CTF parameters were estimated through patch CTF estimation in CryoSPARC. Initial particles were picked and extracted separately using template-based particle picking in CryoSPARC, with preliminary 2D classes as reference. After multiple rounds of 2D classifications and 2D class selections, initial sets of 562,009 and 991,535 particles were selected and refined using non-uniform refinement for each dataset. These particle sets were independently used as a reference for Topaz training, resulting in the extraction of 4,401,976 and 2,826,881 particles from the respective original micrographs through Topaz extraction. Heterogeneous refinement was applied to each dataset, selecting 26.0% and 28.8% of the particles, respectively. These selected particles were then merged into a combined set of 1,939,340 particles. Further heterogeneous refinement of this merged set produced four particle subsets. State V consensus map at 2.21 Å resolution was reconstructed using non-uniform refinement of subset 0 (37.6%, 729,421 particles), while State VI consensus map at 2.24 Å resolution was reconstructed using subset 2 (25.0%, 485,602 particles). For State V, four masks were generated encompassing TBCE, α-tubulin, TBCD, and Arl2-TBCC-RPD. Local refinement was performed for each domain, resulting in four local refined maps. The final map for State V was generated by combining the focused maps with the consensus map using Phenix combine-focused-maps and sharpened with DeepEMhancer. For State VI, TBCD, α-tubulin, TBCD-AD3 and Arl2-TBCC-RPD densities were further refined using local refinement with focused masks for each domain. Particle subsets were selected through 3D classification with focused masks and were subsequently refined using α-tubulin and TBCD-AD3 masks. The final map for State VI was reconstructed by combining the focused maps with the consensus map and sharpened with DeepEMhancer.

3D model building and refinement

The initial models for States I to VI were predicted using AlphaFold3 (*44*), with the version released on 2024/05/23. The models were generated through the AlphaFold server using the default multisequence alignment pipeline and parameters, without alteration. Among the generated models, the model with the highest predicted confidence (model0) was selected as an initial model. The following Uniprot sequences were used as inputs for the prediction: TBCD (Q9BTW9), TBCE (Q15813), TBCC (Q15814), Arl2 (P36404), α-tubulin (Q2XVP4), and β-tubulin (Q767L7).

TBCD–Arl2–β-tub (State I): The atomic model of the TBCD–Arl2–β-tub complex was initially generated using an AlphaFold3-predicted structure. The model was refined to fit the cryo-EM density using Coot and Phenix real-space refinement (*45*, *46*). Subsequently, side-chain conformations and local structural features were manually adjusted using ISOLDE and Coot (*47*). The final atomic model was optimized using Phenix geometry minimization and subsequently manually refined in Coot to improve stereochemistry and model geometry.

TBCDE–Arl2–β-tub (State II): The atomic model of the TBCDE–Arl2–β-tub complex was initially generated using an AlphaFold3-predicted structure. The TBCD–Arl2–β-tub region of the initial model was refined using Coot and Phenix real-space refinement to fit the cryo-EM density. The E-LRR domain was fitted using ChimeraX fit in map command (*43*), followed by real-space refinement. The model was then combined with the TBCD–Arl2–β-tub model in ChimeraX. Subsequently, side-chain conformations and local structural features were manually adjusted using ISOLDE and Coot. The final atomic model was optimized using Phenix geometry minimization and subsequently manually refined in Coot to improve stereochemistry and model geometry.

TBCDE–Arl2$_{GTP}$–β-tub (State III): The initial model was used from the TBCDE–Arl2–β-tub model (State II). From this initial model, the TBCE region was removed, leaving only the UBL domain, and the Arl2 region was roughly aligned with the cryo-EM map. The model was then refined to fit the cryo-EM density using Coot and Phenix real-space refinement. Subsequently, side-chain conformations and local structural features were manually adjusted using ISOLDE and Coot. Finally, the atomic

model was optimized using Phenix geometry minimization and subsequently manually refined in Coot to improve stereochemistry and model geometry.

TBCDE–Arl2$_{GTP}$–αβ-tub (State IV): The atomic model of the TBCDE–Arl2–αβ-tub complex was initially generated using an AlphaFold3-predicted structure. The Arl2 region was initially aligned with the cryo-EM map, followed by refinement of the model to fit the cryo-EM density using Coot and Phenix real-space refinement. Side-chain conformations and local structural features were then manually adjusted using ISOLDE and Coot. Finally, the atomic model was optimized with Phenix geometry minimization and subsequently manually refined in Coot to improve stereochemistry and model geometry.

TBCDEC–Arl2$_{GDP\text{-}AlFx}$–αβ-tub (State V): The atomic model of the TBCDEC–Arl2–αβ-tub complex was initially generated using an AlphaFold3-predicted structure. From this initial model, the TBCC region was removed, leaving only the TBCC-RPD. The model was refined to fit the cryo-EM density using Coot and Phenix real-space refinement. Subsequently, side-chain conformations and local structural features were manually adjusted using ISOLDE and Coot. The final atomic model was optimized using Phenix geometry minimization and subsequently manually refined in Coot to improve stereochemistry and model geometry.

TBCDC–Arl2$_{GDP\text{-}AlFx}$–αβ-tub (State VI): The same initial model was used for State V. From this model, the TBCE region was removed. The model was then refined to fit the cryo-EM density using Coot and Phenix real-space refinement. Side-chain conformations and local structural features were subsequently manually adjusted using ISOLDE and Coot. Finally, the atomic model was optimized using Phenix geometry minimization and subsequently manually refined in Coot to improve stereochemistry and model geometry.

All models are validated by Phenix Comprehensive Validation (cryo-EM) (*48*) and Q-score (*49*).

REFERENCES AND NOTES

1. A. Akhmanova, M. O. Steinmetz, Tracking the ends: A dynamic protein network controls the fate of microtubule tips. *Nat. Rev. Mol. Cell Biol.* **9**, 309–322 (2008). doi: 10.1038/nrm2369; pmid: 18322465
2. C. Janke, M. M. Magiera, The tubulin code and its role in controlling microtubule properties and functions. *Nat. Rev. Mol. Cell Biol.* **21**, 307–326 (2020). doi: 10.1038/s41580-020-0214-3; pmid: 32107477
3. N. B. Gudimchuk, J. R. McIntosh, Regulation of microtubule dynamics, mechanics and function through the growing tip. *Nat. Rev. Mol. Cell Biol.* **22**, 777–795 (2021). doi: 10.1038/s41580-021-00399-x; pmid: 34408299
4. T. Mitchison, M. Kirschner, Dynamic instability of microtubule growth. *Nature* **312**, 237–242 (1984). doi: 10.1038/312237a0; pmid: 6504138
5. T. Horio, H. Hotani, Visualization of the dynamic instability of individual microtubules by dark-field microscopy. *Nature* **321**, 605–607 (1986). doi: 10.1038/321605a0; pmid: 3713844
6. E. Nogales, S. G. Wolf, K. H. Downing, Erratum: Structure of the αβ tubulin dimer by electron crystallography. *Nature* **393**, 191 (1998). doi: 10.1038/30288
7. D. Beckett, G. A. Voth, Unveiling the catalytic mechanism of GTP hydrolysis in microtubules. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2305899120 (2023). doi: 10.1073/pnas.2305899120; pmid: 37364095
8. M. R. Leroux, F. U. Hartl, Protein folding: Versatility of the cytosolic chaperonin TRiC/CCT. *Curr. Biol.* **10**, R260–R264 (2000). doi: 10.1016/S0960-9822(00)00432-2; pmid: 10753735
9. N. J. Cowan, S. A. Lewis, Type II chaperonins, prefoldin, and the tubulin-specific chaperones. *Adv. Protein Chem.* **59**, 73–104 (2001). doi: 10.1016/S0065-3233(01)59003-8; pmid: 11868281
10. D. Gestaut *et al.*, Structural visualization of the tubulin folding pathway directed by human chaperonin TRiC/CCT. *Cell* **186**, 2038 (2023). doi: 10.1016/j.cell.2023.04.004; pmid: 37116473
11. G. Tian *et al.*, Tubulin subunits exist in an activated conformational state generated and maintained by protein cofactors. *J. Cell Biol.* **138**, 821–832 (1997). doi: 10.1083/jcb.138.4.821; pmid: 9265649
12. K. H. Downing, E. Nogales, Tubulin structure: Insights into microtubule properties and functions. *Curr. Opin. Struct. Biol.* **8**, 785–791 (1998). doi: 10.1016/S0959-440X(98)80099-7; pmid: 9914260
13. G. Tian, N. J. Cowan, Tubulin-specific chaperones: Components of a molecular machine that assembles the α/β heterodimer. *Methods Cell Biol.* **115**, 155–171 (2013). doi: 10.1016/B978-0-12-407757-7.00011-6; pmid: 23973072
14. S. Nithianantham *et al.*, Tubulin cofactors and Arl2 are cage-like chaperones that regulate the soluble αβ-tubulin pool for microtubule dynamics. *eLife* **4**, e08811 (2015). doi: 10.7554/eLife.08811; pmid: 26208336
15. J. D. Sharer, J. F. Shern, H. Van Valkenburgh, D. C. Wallace, R. A. Kahn, ARL2 and BART enter mitochondria and bind the adenine nucleotide transporter. *Mol. Biol. Cell* **13**, 71–83 (2002). doi: 10.1091/mbc.01-05-0245; pmid: 11809823
16. L. E. Newman *et al.*, The ARL2 GTPase is required for mitochondrial morphology, motility, and maintenance of ATP levels. *PLOS ONE* **9**, e99270 (2014). doi: 10.1371/journal.pone.0099270; pmid: 24911211
17. Y. Gao, I. E. Vainberg, R. L. Chow, N. J. Cowan, Two cofactors and cytoplasmic chaperonin are required for the folding of alpha- and beta-tubulin. *Mol. Cell. Biol.* **13**, 2478–2485 (1993). doi: 10.1128/mcb.13.4.2478-2485.1993; pmid: 8096061
18. M. Lopez-Fanarraga, J. Avila, A. Guasch, M. Coll, J. C. Zabala, Review: Postchaperonin tubulin folding cofactors and their role in microtubule dynamics. *J. Struct. Biol.* **135**, 219–229 (2001). doi: 10.1006/jsbi.2001.4386; pmid: 11580271
19. D. Kortazar *et al.*, Role of cofactors B (TBCB) and E (TBCE) in tubulin heterodimer dissociation. *Exp. Cell Res.* **313**, 425–436 (2007). doi: 10.1016/j.yexcr.2006.09.002; pmid: 17184771
20. J. Al-Bassam, Revisiting the tubulin cofactors and Arl2 in the regulation of soluble αβ-tubulin pools and their effect on microtubule dynamics. *Mol. Biol. Cell* **28**, 359–363 (2017). doi: 10.1091/mbc.e15-10-0694; pmid: 28137948
21. T. J. Mitchison, Localization of an exchangeable GTP binding site at the plus end of microtubules. *Science* **261**, 1044–1047 (1993). doi: 10.1126/science.8102497; pmid: 8102497
22. J. W. Francis, L. E. Newman, L. A. Cunningham, R. A. Kahn, A Trimer Consisting of the Tubulin-specific Chaperone D (TBCD), Regulatory GTPase ARL2, and β-Tubulin Is Required for Maintaining the Microtubule Network. *J. Biol. Chem.* **292**, 4336–4349 (2017). doi: 10.1074/jbc.M116.770909; pmid: 28126905
23. J. W. Francis *et al.*, Nucleotide Binding to ARL2 in the TBCD·ARL2·β-Tubulin Complex Drives Conformational Changes in β-Tubulin. *J. Mol. Biol.* **429**, 3696–3716 (2017). doi: 10.1016/j.jmb.2017.09.016; pmid: 28970104
24. M. Serna *et al.*, The structure of the complex between α-tubulin, TBCE and TBCB reveals a tubulin dimer dissociation mechanism. *J. Cell Sci.* **128**, 1824–1834 (2015). pmid: 25908846
25. G. Tian, A. Bhamidipati, N. J. Cowan, S. A. Lewis, Tubulin folding cofactors as GTPase-activating proteins. GTP hydrolysis and the assembly of the alpha/beta-tubulin heterodimer. *J. Biol. Chem.* **274**, 24054–24058 (1999). doi: 10.1074/jbc.274.34.24054; pmid: 10446175
26. A. Taheri, Z. Wang, B. Singal, F. Guo, J. Al-Bassam, Cryo-EM structures of the tubulin cofactors reveal the molecular basis for the biogenesis of alpha/beta-tubulin. bioRxiv 2024.01.29.577855 [Preprint] (2024); https://doi.org/10.1101/2024.01.29.577855.
27. M. Hanzal-Bayer, L. Renault, P. Roversi, A. Wittinghofer, R. C. Hillig, The complex of Arl2-GTP and PDE delta: From structure to function. *EMBO J.* **21**, 2095–2106 (2002). doi: 10.1093/emboj/21.9.2095; pmid: 11980706
28. D. Kortazar *et al.*, Native tubulin-folding cofactor E purified from baculovirus-infected Sf9 cells dissociates tubulin dimers. *Protein Expr. Purif.* **49**, 196–202 (2006). doi: 10.1016/j.pep.2006.03.005; pmid: 16624573
29. A. Bhamidipati, S. A. Lewis, N. J. Cowan, ADP ribosylation factor-like protein 2 (Arl2) regulates the interaction of tubulin-folding cofactor D with native tubulin. *J. Cell Biol.* **149**, 1087–1096 (2000). doi: 10.1083/jcb.149.5.1087; pmid: 10831612
30. F. Bartolini *et al.*, Functional overlap between retinitis pigmentosa 2 protein and the tubulin-specific chaperone cofactor C. *J. Biol. Chem.* **277**, 14629–14634 (2002). doi: 10.1074/jbc.M200128200; pmid: 11847227
31. A. Shemesh *et al.*, Effect of tubulin self-association on GTP hydrolysis and nucleotide exchange reactions. *Biochim. Biophys. Acta. Proteins Proteom.* **1871**, 140869 (2023). doi: 10.1016/j.bbapap.2022.140869; pmid: 36400388
32. R. Zhang, B. LaFrance, E. Nogales, Separating the effects of nucleotide and EB binding on microtubule structure. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E6191–E6200 (2018). doi: 10.1073/pnas.1802637115; pmid: 29915050
33. C. Akıl *et al.*, Structure and dynamics of Odinarchaeota tubulin and the implications for eukaryotic microtubule evolution. *Sci. Adv.* **8**, eabm2225 (2022). doi: 10.1126/sciadv.abm2225; pmid: 35333570
34. G. Tian *et al.*, Pathway leading to correctly folded beta-tubulin. *Cell* **86**, 287–296 (1996). doi: 10.1016/S0092-8674(00)80100-2; pmid: 8706133
35. S. Chakraborty *et al.*, Cryo-ET suggests tubulin chaperones form a subset of microtubule lumenal particles with a role in maintaining neuronal microtubules. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2404017121 (2025). doi: 10.1073/pnas.2404017121; pmid: 39888918
36. J. Zhou *et al.*, Structural insights into the mechanism of GTP initiation of microtubule assembly. *Nat. Commun.* **14**, 5980 (2023). doi: 10.1038/s41467-023-41615-w; pmid: 37749104
37. M. A. Hoyt, J. P. Macke, B. T. Roberts, J. R. Geiser, Saccharomyces cerevisiae PAC2 functions with CIN1, 2 and 4 in a pathway leading to normal microtubule stability. *Genetics* **146**, 849–857 (1997). doi: 10.1093/genetics/146.3.849; pmid: 9215891

38. P. A. Radcliffe, L. Vardy, T. Toda, A conserved small GTP-binding protein Alp41 is essential for the cofactor-dependent biogenesis of microtubules in fission yeast. *FEBS Lett.* **468**, 84–88 (2000). doi: 10.1016/S0014-5793(00)01202-3; pmid: 10683446
39. S. Kofuji, A. T. Sasaki, GTP metabolic reprogramming by IMPDH2: Unlocking cancer cells' fuelling mechanism. *J. Biochem.* **168**, 319–328 (2020). doi: 10.1093/jb/mvaa085; pmid: 32702086
40. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017). doi: 10.1038/nmeth.4169; pmid: 28165473
41. A. Punjani, H. Zhang, D. J. Fleet, Non-uniform refinement: Adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020). doi: 10.1038/s41592-020-00990-8; pmid: 33257830
42. R. Sanchez-Garcia *et al.*, DeepEMhancer: A deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021). doi: 10.1038/s42003-021-02399-1; pmid: 34267316
43. E. F. Pettersen *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021). doi: 10.1002/pro.3943; pmid: 32881101
44. J. Abramson *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024). doi: 10.1038/s41586-024-07487-w; pmid: 38718835
45. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010). doi: 10.1107/S0907444910007493; pmid: 20383002
46. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019). doi: 10.1107/S2059798319011471; pmid: 31588918
47. T. I. Croll, ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530 (2018). doi: 10.1107/S2059798318002425; pmid: 29872003
48. P. V. Afonine *et al.*, New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D Struct. Biol.* **74**, 814–840 (2018). doi: 10.1107/S2059798318009324; pmid: 30198894
49. G. Pintilie *et al.*, Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat. Methods* **17**, 328–334 (2020). doi: 10.1038/s41592-020-0731-1; pmid: 32042190

ACKNOWLEDGMENTS

The cryo-EM data were collected at the Center for Macromolecular and Cell Imaging (CMCI) , the National Center for Inter-University Research Facilities (NCIRF), and the Institute of Membrane Proteins (IMP) and processed at the Global Science Experimental Data Hub Center (GSDC) at Korea Institute of Science facilities. We thank G. Pintilie (SLAC National Accelerator Laboratory, Stanford University) for his assistance in generating molecular movies used in this study. **Funding:** This work was supported by the Korean National Research Foundation (RS-2021-NR056571, RS-2020-NR049538, RS-2020-NF000307, RS-2024-00344154, RS-2024-00440289, and RS-2025-00559184) and Suh Kyungbae Foundation (SUHF) to S.-H.R. **Author contributions:** Y.S. and H.K. performed cryo-EM imaging, data collection, and structural analysis/model building. K.B. and Y. -W. P. performed cloning, protein purification, and biochemical experiments. S.-H.R. conceptualized and designed the project and wrote the manuscript in collaboration with the other authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The 3D cryo-EM density maps have been deposited in the Electron Microscopy Data Bank under accession codes EMD-63572 (TBCD–Arl2–β-tub, State I), EMD-63573 (TBCDE–Arl2–β-tub, State II), EMD-63574 (TBCDE–$Arl2_{GTP}$–β-tub, State III), EMD-63575 (TBCDE–$Arl2_{GTP}$–αβ-tub, State IV), EMD-63576 (TBCDEC–$Arl2_{GDP\text{-}AlFx}$–αβ-tub, State V) and EMD-63577 (TBCDC–$Arl2_{GDP\text{-}AlFx}$–αβ-tub, State VI). Coordinates for the 3D cryo-EM density maps have been deposited in the Protein Data Bank under accession codes PDB 00009M1I, 00009M1J, 00009M1K, 00009M1L, 00009M1M, and 00009M1N, respectively. Consensus maps and focused maps have been deposited in the Electron Microscopy Data Bank under accession code EMD-63569 (State II consensus map), EMD-63570 (State II TBCE-LRR domain–focused map), EMD-63376 (State III consensus map), EMD-63377 (State III Arl2 and TBCD-AD1–focused map), EMD-63378 (State III TBCD-AD3 and TBCE-UBD–focused map), EMD-63546 (State IV consensus map), EMD-63547 (State IV TBCE and α-tubulin–focused map), EMD-63548 (State IV Arl2 and TBCD-AD1–focused map), EMD-63549 (State V consensus map), EMD-63550 (State V TBCE–focused map), EMD-63551 (State V TBCE and α-tubulin–focused map), EMD-63552 (State V TBCC-RPD and Arl2–focused map), EMD-63553 (State V TBCD and Arl2–focused map), EMD-63554 (State VI consensus map), EMD-63555 (State VI α-tubulin–focused map), EMD-63556 (State VI TBCD and Arl2–focused map), EMD-63557 (State VI TBCC-RPD and Arl2–focused map), and EMD-63558 (State VI TBCD-AD3–focused map). All other data needed to support the conclusions in this manuscript can be found in the main text or supplementary materials. **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ady2708
Figs. S1 to S7; Table S1; MDAR Reproducibility Checklist; Movies S1 and S2

Submitted 16 April 2025; accepted 25 July 2025

10.1126/science.ady2708

RESEARCH ARTICLE SUMMARY

GREENHOUSE GASES

Global methane action pays for itself at least six times over

Thomas Stoerk*, James Rising, Drew Shindell, Simon Dietz

Full article and list of author affiliations: https://doi.org/10.1126/science.adu7392

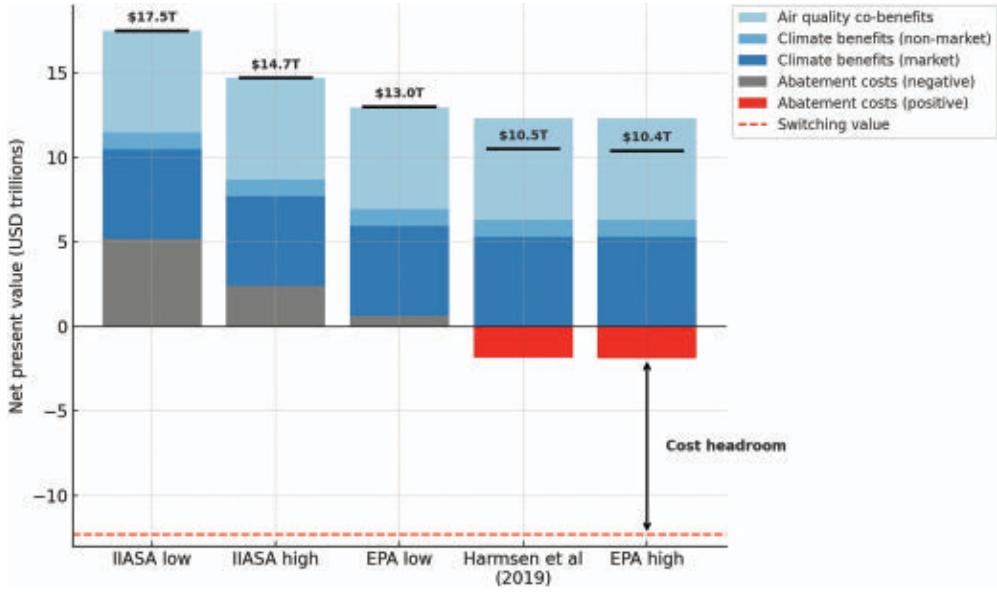
INTRODUCTION: Methane is a powerful but short-lived greenhouse gas, responsible for about one-third of the warming caused by well-mixed greenhouse gases since the 19th century. Its concentration in the atmosphere is once again rising rapidly owing to emissions from fossil fuels, agriculture, and waste, as well as anthropogenic and natural emissions from wetlands. Because of these two factors, reducing methane emissions has been identified as one of the most effective ways to curb near-term global warming. In response, the Global Methane Pledge, launched at COP26 in 2021, set an international target to reduce anthropogenic methane emissions by at least 30% from 2020 levels by 2030. Understanding the economic case for such action is crucial, especially given concerns about costs and the distribution of climate impacts across countries.

RATIONALE: We provide a comprehensive economic assessment of global methane abatement, quantifying both benefits and costs. Our study improves on prior work by incorporating recent climate modeling, country-level damage estimates, and the effects of climate tipping points. We anchor our analysis on methane scenarios consistent with the Global Methane Pledge, using an integrated assessment model (IAM) to capture the impact of methane abatement on global temperatures, economic damages, and key tipping elements such as ice sheets and ocean currents. In addition to modeling climate-related benefits, we include improved air quality from reduced methane-driven ozone, and we evaluate abatement costs using detailed sectoral estimates from the Global Methane Assessment.

RESULTS: Our findings show that global methane abatement yields large net economic benefits. Achieving the Global Methane Pledge results in more than $1 trillion in annual avoided market damages by 2050. Benefits are greater in lower-income countries because of their higher vulnerability to climate change. Methane action also reduces climate risk, including the likelihood and severity of extreme damage outcomes (climate tail risk). It lowers the intensity of several key tipping points: For example, by 2050, the likelihood of Amazon rainforest dieback drops by 8% and Indian monsoon disruption by 13%. Even using conservative assumptions about costs, the benefit-cost ratio of methane action is at least three—rising to more than six when health cobenefits are included. Some cost estimates suggest that methane mitigation could even be profitable owing to energy recovery and other gains. Finally, we provide estimates of the "social cost of methane" [the monetary value of future damages from 1 metric ton of methane emissions (tCH_4)], finding a global average of $7381/$tCH_4$, with higher domestic values for economies such as the US, EU, China, and India—implying strong economic self-interest in methane reductions.

CONCLUSION: Our analysis shows that global methane mitigation is not only feasible but also economically compelling. The benefits—spanning avoided climate damages and improved air quality—consistently outweigh the costs under a wide range of scenarios. The gains are especially pronounced for low- and middle-income countries, underscoring the policy's equity-enhancing potential. Given that major economies also stand to gain considerably on their own, global cooperation is not strictly necessary for substantial methane action to be in their national interest. □

*Corresponding author. Email: t.a.stoerk@lse.ac.uk Cite this article as T. Stoerk *et al.*, *Science* **390**, eadu7392 (2025). DOI: 10.1126/science.adu7392



Net economic benefit of global methane action from 2020 to 2050. Net present value of climate benefits (market and nonmarket), air-quality cobenefits, and abatement costs under five methane abatement cost scenarios. Horizontal black lines show net benefits. The dashed line indicates the switching value threshold and available cost headroom before net present value is reduced to zero. The discount rate is 2.5%. IIASA, International Institute for Applied Systems Analysis; EPA, US Environmental Protection Agency; Harmsen *et al.*, M. Harmsen *et al.*, *Environ. Sci. Policy* **99**, 136–149 (2019).

GREENHOUSE GASES

Global methane action pays for itself at least six times over

Thomas Stoerk[1,2]*, James Rising[3], Drew Shindell[4], Simon Dietz[2]

We provide a comprehensive assessment of the economic benefits and costs of global methane emissions abatement, anchored on the Global Methane Pledge. We use an integrated assessment model to estimate avoided climate damages at the global and country levels, including quantification of tipping points and risk. We further estimate air-quality cobenefits and methane abatement costs. We find that global methane action would be highly beneficial, yielding a benefit-cost ratio of at least six. It would provide larger benefits in lower-income countries, and it would reduce tipping point intensity and risk. We provide estimates of the social cost of methane to compare with previous literature and show that these estimates imply that key economies, such as the United States, European Union, and China, should be self-interested to abate methane emissions substantially.

Methane is a potent short-lived climate pollutant (SLCP) that has contributed around one-third of the total warming from well-mixed greenhouse gases (GHGs) since the 19th century (*1*). Having plateaued in the early 2000s, atmospheric concentrations of methane are once again increasing at a rapid rate because of anthropogenic emissions from fossil fuel production and distribution, agriculture, and waste, as well as anthropogenic and natural emissions from wetlands (*2*–*6*).

Because it is a major driver of current warming and its atmospheric concentration responds rapidly to emissions reductions, methane abatement has been identified as "very likely to be the most powerful lever in reducing near-term warming" (*7*). Moreover, substantial technical potential has been identified to reduce methane emissions, much of which is argued to be negative cost or low cost (*7*–*11*).

Previous work has argued that climate policy is too focused on long-term climate stability (*11*, *12*)—for example, on mid-century, net-zero CO_2 emissions as the key target to be adopted by both state and nonstate actors. Even the more near-term Nationally Determined Contributions to the 2015 UN Paris Agreement have been found to contain variable, but generally low, coverage of methane-focused mitigation measures (*13*). In the background to policy-making, the accounting practice of converting methane into CO_2 equivalents by using its 100-year global warming potential (GWP) has been argued to contribute to overlooking the importance of methane because such a conversion understates the impact of methane in the short term (*12*). The contrary view is that methane abatement, alongside mitigation of other SLCPs, risks delaying efforts to abate CO_2 and can be "deferred without much harm" if the ultimate goal is controlling long-term warming (*14*).

Policy-makers appear to be listening to arguments in favor of methane action. The Global Methane Pledge, launched at COP26 in 2021, commits participating countries to take voluntary, domestic actions in service of at least a 30% reduction in global anthropogenic methane emissions below the 2020 level by 2030 (*15*). At the time of writing, 159 countries have signed on (including the EU and US but not China, India, or Russia), and some have reflected the global ambition of the Pledge in domestic targets and plans (*13*).

Understanding the role of methane abatement in climate policy can be aided by quantifying the benefits of methane abatement and comparing them with the costs. Previous studies have quantified the benefits of global methane abatement scenarios in terms of physical climate variables such as carbon budgets (*16*, *17*), global mean temperatures (*9*, *11*, *18*–*23*), and sea levels (*24*). The 2021 Global Methane Assessment extended the literature by quantifying selected economic and health benefits of global methane abatement, including heat- and ozone-related mortality, ozone-related morbidity, lost labor productivity owing to extreme heat, and crop yield losses via both climate and ozone effects (*7*). (*23*) have provided a simple calculation of the overall economic benefits of methane abatement by feeding global mean temperature pathways into an aggregate climate damage function.

In addition to these studies on the benefits of global methane action scenarios, there is a literature on the social cost of methane (SC-CH_4)—i.e., the discounted present value of the future stream of damages from an incremental (marginal) emission of the gas (*25*–*35*). The literature on the SC-CH_4 is much smaller than its counterpart on the social cost of carbon dioxide (SC-CO_2), but it is valuable because social (or marginal) costs provide a natural and consistent way of comparing GHGs with different atmospheric lifetimes (*36*). As (*19*) pointed out, "direct comparisons of the climate influence of SLCPs and CO_2 require making a judgment about the relative importance of short and long time scales." Economics would compare the SC-CH_4 with the SC-CO_2, as doing so does not depend on an arbitrary choice of time horizon, unlike GWP (e.g., GWP20 versus GWP100), although the discount rate becomes a key choice (*37*). The SC-CH_4 is also an important policy quantity given its routine application in cost-benefit analyses (CBAs) of specific methane abatement policies (e.g., rules on methane flaring in oil and gas). However, the standard practice in national policy analysis of multiplying GHG emissions reductions by the social cost of a GHG (SC-GHG) cannot, in general, be used to quantify large global emissions reductions brought about by agreements such as the Global Methane Pledge. That is because the marginal approximation that is relied on can be highly inaccurate for large, nonmarginal changes in emissions (*38*, *39*). Therefore, an assessment of the total economic benefits of global methane action should be based on running emissions scenarios through an integrated assessment model (IAM).

In this paper, we seek to advance the literature by providing a comprehensive assessment of the economic benefits of global methane abatement, anchored on the Global Methane Pledge, and comparing them with the costs. This includes a distributional analysis across countries and an analysis of how global methane abatement affects tipping points and macroeconomic risk. We proceed in three steps. First, we create a set of emissions scenarios, which simulate methane action at the global level against a background of emissions of CO_2 and other GHGs and radiative forcing agents. These scenarios are anchored on successful achievement in 2030 of the Global Methane Pledge, followed by continued methane abatement post-2030.

The second step is to run the scenarios through the Model for Economic Tipping point Analysis (META) IAM (*40*, *41*) to estimate total economic benefits. META has several useful features for present purposes: (i) it incorporates a recent Simple Climate Model from the climate science literature (*42*), thus producing global mean temperature projections consistent with observations and Earth System Models; (ii) it provides relatively comprehensive benefit/damage estimates, including both market and nonmarket damages, via temperature and sea level mechanisms; (iii) it estimates damages at the country level for 179 countries, thus enabling the distributional effects of emissions abatement to be analyzed; (iv) it can quantify the role of climate tipping points (including methane-based tipping points) and climate tail risk in the overall benefits of methane action (however,

[1]Economics and Research Department, National Bank of Belgium, Brussels, Belgium. [2]Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science, London, UK. [3]School of Marine Science and Policy, University of Delaware, Newark, DE, USA. [4]Nicholas School of the Environment, Duke University, Durham, NC, USA. ***Corresponding author. Email: t.a.stoerk@lse.ac.uk**

the coverage of tipping point damage mechanisms is incomplete, a reflection of the underlying literature).

The third step brings in methane abatement costs and air-quality cobenefits. For the former, we aggregate sectoral estimates of methane abatement costs from (*43*) and (*10*), which were used in the Global Methane Assessment (*7*), as well as using the methane marginal abatement cost (MAC) curve estimated by (*8*). For the latter, we use data on premature deaths due to respiratory and cardiovascular diseases and on the value of a statistical life (VSL), also from the Global Methane Assessment (*7*).

In addition to quantifying the total benefits and costs of global methane action, we also include an analysis of the SC-CH_4. This facilitates a comparison with previous literature, which has tended to be about marginal rather than total damages. In addition, with META, we can provide a set of country-level estimates of the SC-CH_4. In light of current geopolitics, this enables us to evaluate incentives of countries to act in their own self-interest rather than just in pursuit of a global agreement such as the Global Methane Pledge.

Results

Methane emissions and effect on temperatures

Figure 1 plots methane emissions and resulting global mean temperatures under our methane action scenarios: central, low emissions, and high emissions. The central scenario contains a subset of 34 of the 1.5°C scenarios from the scenario database that approximate the Global Methane Pledge's 30% reduction. The low- and high-emissions scenarios are subsets of the 1.5°C scenarios that bracket the central scenario and give, respectively, ~50% bigger and smaller cuts than the central scenario. The scenario spread is intended to reflect uncertainty about technical abatement potential across IAMs and about policy implementation (within the context of methane action above current trends). Our preferred specification chooses a background shared socioeconomic pathway (SSP) scenario of SSP2-4.5, which we take to be the SSP scenario corresponding most closely to a current policies or trends storyline (*44*). This sets both baseline methane emissions and emissions of other GHGs and forcing agents. In fig. S1 in the supplementary materials, we also show a high-emissions background scenario (SSP3-7.0) and a 1.5°C background scenario. The former amounts to a storyline in which the Global Methane Pledge is successfully implemented but the rest of the world's climate agenda fails completely. This is a less likely combination in our view, although observations imply methane emissions specifically are currently increasing faster than projected on SSP2-4.5 (*45*). The use of a 1.5°C background scenario just assesses whether the Global Methane Pledge requires similar, more, or less methane abatement than the full suite of 1.5°C scenarios.

The central methane action scenario reduces methane emissions by 30% in 2030 relative to 2020 (implementation of the Global Methane Pledge), 47% in 2050, and 62% in 2100 (Fig. 1, left). The methane action high-low scenario range is a 19 to 46% cut in 2030, 34 to 63% in 2050, and 52 to 72% in 2100, all relative to 2020. The supplementary materials show that although the Global Methane Pledge is well within the range of 1.5°C scenarios, it is less ambitious than the mean cut in 2030 across 1.5°C scenarios. Emissions are 4% higher under the Global Methane Pledge than in the mean 1.5°C scenario.

The central methane action scenario reduces global mean surface temperature change by an expected 0.02°C in 2030, 0.13°C in 2050, and 0.18°C in 2100 (Fig. 1, right). The temperature reduction under the methane action scenarios flattens out after 2070 because baseline methane emissions reductions under SSP2-4.5 follow a similar trajectory to the methane action scenarios, albeit at an overall higher emissions level. The range of uncertainty in the expected temperature response to the high-low methane scenario variants is 0.09° to 0.19°C in 2050. Temperature uncertainty conditional on the central methane scenario is also 0.09° to 0.19°C in 2050. This comes from parametric uncertainty in the Finite Amplitude Impulse Response (FaIR) Simple Climate Model, as well as tipping points in META that affect temperature, and is quantified as the 90% confidence interval. The expected 0.13°C of avoided warming in 2050 on the central methane scenario is lower than the ~0.3°C of avoided warming reported in the Global Methane Assessment (*7*) on the basis of similar emissions-temperature impulse-response functions to those in FaIR. The difference is presumably due to higher baseline emissions in the Global Methane Assessment, more consistent with no policy than current policies. Figure S2 in the supplementary materials shows that the temperature reduction is 0.2°C in 2050 relative to an SSP3-7.0 baseline. This is still lower than that in the Global Methane Assessment, but both scenario and temperature-response uncertainty extend to ~0.3°C.

Total economic benefits of methane action

Figure 2 plots the total economic benefits (avoided market damages) of the methane action scenarios over the rest of this century. Under the SSP2-4.5 background scenario, climate change causes mean damages

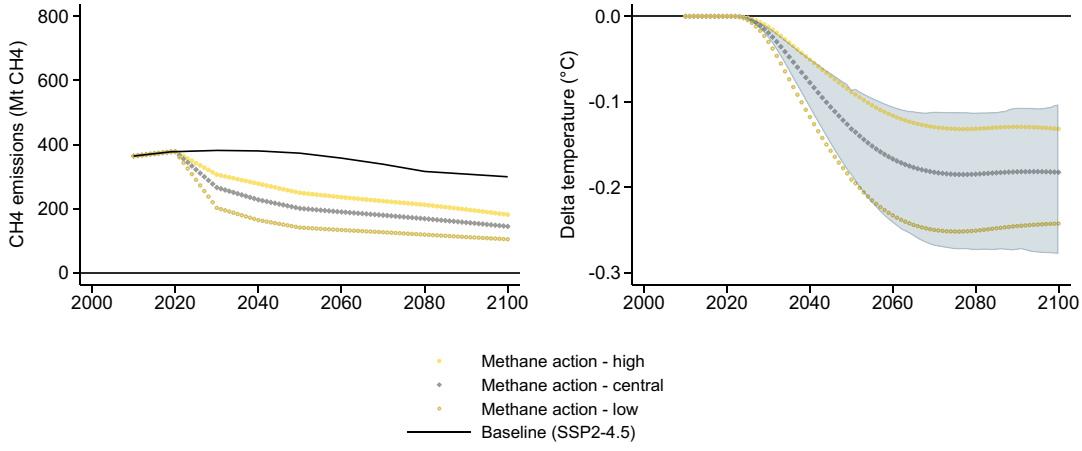


Fig. 1. Methane emissions and effect on global mean temperature of methane action scenarios. (**Left**) Methane emissions. (**Right**) Effect on global mean temperature. Background emissions scenario is SSP2-4.5. The shaded area is the 90% confidence interval of temperature outcomes pertaining to the central methane action scenario, based on parametric uncertainty in META sampled by 1000 Monte Carlo runs. "Methane action – high" and "Methane action – low" describe methane action scenarios resulting in high and low emissions, respectively. Specification includes tipping points.

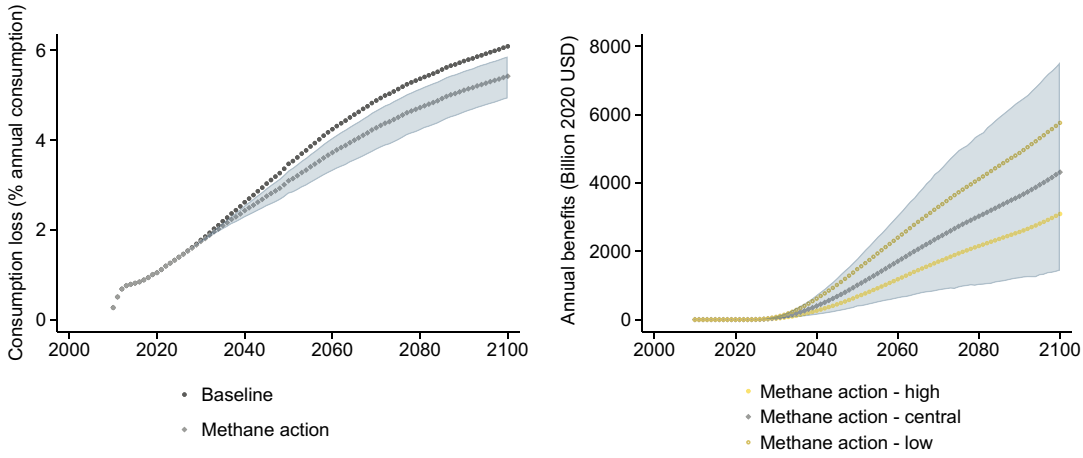


Fig. 2. Reduction in total climate damages due to methane action. (**Left**) Total climate damages expressed as a percentage loss in consumption relative to a no-climate-change counterfactual. (**Right**) Gross monetary value of avoided damages (i.e., benefits of methane action). The shaded area is the 90% confidence interval, based on parametric uncertainty in META sampled by 1000 Monte Carlo runs. Specification includes market damages and tipping points.

equivalent to a 1.8% loss in global consumption in 2030, rising to 3.5% in 2050 and 6.1% in 2100 [because savings rates are fixed in META, these percentages also apply to global gross domestic product (GDP)]. These are relative to counterfactual consumption along SSP2-4.5 without climate change. Focusing on 2050, the central methane action scenario reduces damages by 0.4 percentage points, from 3.5% of global consumption to 3.1%. In absolute money terms, this reduction in damages—the gross benefit of global methane action—equates to a little more than $1 trillion/year in 2050 (in 2020 prices). The range of uncertainty across the high-low methane scenario variants is a reduction in damages of 0.3 to 0.6 percentage points. The range of uncertainty (90% confidence interval) conditional on the central methane scenario is larger than this, running from 0.2 to 0.7 percentage points. This comes from the joint effect of many parametric uncertainties in META, including climate uncertainties in FaIR, uncertainties in the tipping point modules and in the interactions between tipping points, and uncertainties in the temperature and sea level damage functions. Further analysis in fig. S3 in the supplementary materials shows that, as expected, total climate damages are higher under the SSP3-7.0 background scenario and that the reduction in damages due to methane action is around a quarter higher than under SSP2-4.5 (a reduction in damages of 0.5 percentage points in 2050).

The above estimates of total economic benefits can only be properly understood when compared with the total costs of methane action. Accounting for nonmarket climate damages and the direct health co-benefits of methane action via improved air quality is also important. META contains a nonmarket damages module, but, as a social-cost IAM focused on damages (*46*), META does not include abatement costs or air-quality cobenefits. Therefore, we bring in data on abatement costs and air-quality cobenefits from previous major studies (*7*, *8*) (materials and methods). Table 1 reports the results of the CBA. Three of five estimates of abatement costs are negative throughout; thus, the central methane action scenario would pay for itself irrespective of climate benefits. The mechanism behind negative abatement costs is the opportunity for private profits, including revenue from methane capture and use in the energy sector, reduced disposal costs and increased revenue from waste recycling instead of landfilling, and yield increases from abatement measures in rice farming. In these three cost scenarios [International Institute for Applied Systems Analysis (IIASA) low, IIASA high, and US Environmental Protection Agency (EPA) low], the net present value (NPV) of the central methane scenario ranges from $0.6 trillion to $5.2 trillion without climate benefits, $7.0 trillion to $11.5 trillion including climate benefits, and $13.0 trillion to $17.5 trillion additionally including air-quality cobenefits (2.5% discount rate).

Table 1. Comparison of benefits and costs of central methane scenario. All values are billions of 2020 US dollars. Discount rate is 2.5%. Values for 2030, 2040, and 2050 are snapshots of benefits and abatement costs in single years. NPVs represent the discounted sums for years 2020 through 2050.

Components of benefits and costs		2030	2040	2050	NPV
Climate benefits (market)		52.2	401.0	1008.5	5310.0
Climate benefits (nonmarket)		9.1	70.3	200.4	1001.8
Air-quality cobenefits		298.8	393.3	444.6	6015.7
Abatement costs	EPA low	−34.0	−39.7	−45.3	−638.3
	EPA high	108.2	113.8	119.5	1885.2
	IIASA low	−226.6	−342.5	−458.4	−5176.3
	IIASA high	−186.4	−123.1	−59.7	−2387.7
	Harmsen *et al*. (*8*)	18.3	130.9	437.6	1847.8
Discounted net climate benefits (market and nonmarket)	EPA low	74.4	311.8	597.9	6950.1
	EPA high	−36.6	218.2	519.4	4426.6
	IIASA low	224.9	496.6	794.8	11,488.1
	IIASA high	193.5	362.7	604.8	8699.4
	Harmsen *et al*. (*8*)	33.6	207.7	367.7	4464.0
Discounted net climate benefits (market and nonmarket) + air-quality cobenefits	EPA low	307.8	551.8	809.9	12,965.7
	EPA high	196.7	458.2	731.4	10,442.2
	IIASA low	458.3	736.6	1006.8	17,503.7
	IIASA high	426.9	602.7	816.8	14,715.1
	Harmsen *et al*. (*8*)	267.0	447.7	579.7	10,479.6

Although "bottom-up" estimates of GHG abatement costs can yield substantial negative-cost potential (*47*), such measures face behavioral, financial, informational, and structural barriers, which prevent them from being realized. These can be conceptualized as hidden costs of adoption. Especially the low-cost scenarios implicitly assume that implementation of the Global Methane Pledge reduces the barriers to adoption of negative-cost measures. Alternatively, a conservative approach would be to place particular emphasis on the two scenarios with positive abatement costs, EPA high and Harmsen *et al.* (*8*). The

latter scenario explicitly assumes there are no negative-cost measures. However, again, the benefits of methane action far exceed the costs, giving an NPV of $4.4 trillion to $4.5 trillion ($10.4 trillion to $10.5 trillion including air-quality cobenefits), or equivalently, a benefit-cost ratio of three (six including air-quality cobenefits). As a further alternative, we can back-calculate what constant stream of undiscounted abatement costs would be necessary to drive the NPV of the central methane scenario to zero. This "switching value" is $301.6 billion/year ($589.0 billion/year including air-quality cobenefits), much higher than any of the cost scenarios except for Harmsen *et al.* (*8*) toward 2050, ignoring air-quality cobenefits. The cost-benefit comparison is also conservative in the sense that we only consider costs and benefits in the period 2020 to 2050. Methane has an atmospheric residence time of ~9 years, so this approach omits benefits of abatement toward 2050 that occur after the end of the horizon. Persistence of climate damages has a similar effect.

Distributional effects of methane action

Climate damages are heterogeneous across the world for reasons of differential exposure and sensitivity or adaptive capacity. META incorporates both of these sources of heterogeneity. Insofar as these are correlated with national income, so will be the benefits of methane action. Figure 3 explores the relationship between national income and benefits of methane action directly by aggregating country-level avoided market damages from methane action in accordance with the World Bank's four country groupings: low income, lower-middle income, upper-middle income, and high income. These damages are differences in consumption levels before utility or welfare valuation. Larger benefits from methane action are concentrated in low-, lower-middle–, and to a lesser extent, upper-middle–income countries, whereas high-income countries, which tend to be located at higher latitudes, benefit the least. With the exception of low-income versus lower-middle–income countries, group differences across countries are statistically significant at the 10% level or lower for any pairwise comparison of World Bank income groups. Figure S4 in the supplementary materials repeats this analysis under the SSP3-7.0 background scenario, showing that the equity benefits of methane mitigation are robust to switching to an emissions scenario with substantially higher emissions. Figure S5 in the supplementary materials also repeats this analysis on a scenario with rapid methane action plus action on CO_2 to explore whether the effects of methane abatement on inequality are different from CO_2 abatement. We find the effects are very similar, which indicates that action on methane and CO_2 are equally pro-poor. Lastly, fig. S6 in the supplementary materials repeats the analysis for avoided market and nonmarket damages together. It shows that, again, benefits are negatively correlated with income, but the differences are now smaller and statistically insignificant because nonmarket benefits increase with income.

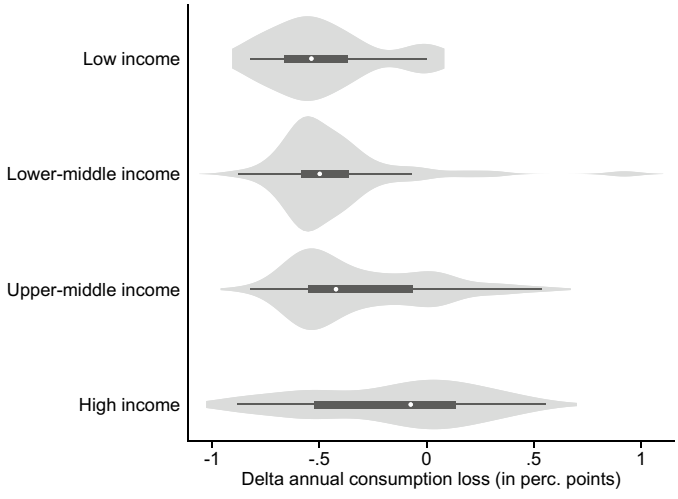


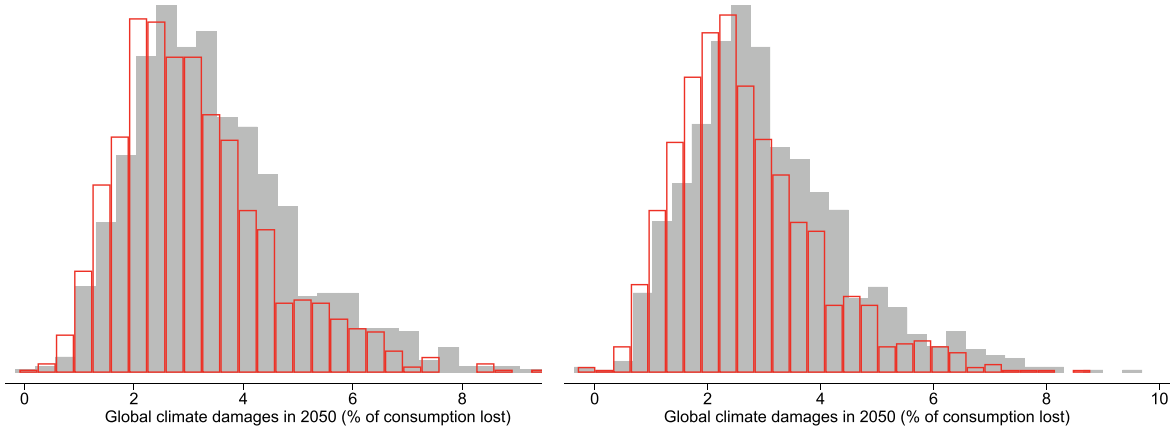
Fig. 3. Change in climate damages in 2050 by central methane action scenario, disaggregated by World Bank country income group. Violin plots combine a kernel density plot and a box plot. Box plots contain estimates for the median (white dot), interquartile range (black box), and ±1.5 times the interquartile range (range of thin dark gray line). Negative values signify avoided climate change damages. Specification includes market damages and tipping points.

Effect of methane action on climate risk

Figure 4 reports the effect of methane action on climate risk. Methane action reduces climate risks, shifting the probability density function of damages in 2050 to the left (downward). Mean damages (market) fall from 3.5 to 3.1% as mentioned above, and the standard deviation of damages decreases from 1.5 to 1.4. As measures of tail risk, we compute value at risk (VaR) and conditional VaR (CVaR) at the 95% and 99% confidence levels. These are defined as damages at the 95th/99th percentiles (VaR) and average damages above the 95th/99th percentiles (CVar). Methane action reduces VaR at the 95% confidence level from 6.4 to 5.9% and at the 99% confidence level from 7.8 to 7.3%. CVaR at the 95% confidence level is reduced from 7.4 to 6.7% and at the 99% confidence level from 8.9 to 8.0%. Thus, methane action reduces climate risk in general, including tail risk. Without tipping points, the effects of methane action are very similar or slightly larger in relative terms, which indicates that methane action is slightly more effective in reducing damages from other channels. Additional analysis in fig. S7 in the supplementary materials shows that methane action similarly reduces climate risk under a SSP3-7.0 background scenario, albeit on higher baseline climate damages.

This raises the question, what is the effect of methane action on tipping points themselves? As explained in more detail in (*40*) and in the supplementary materials, economic models of tipping points synthesized by META belong to two distinct classes. One class is process-based models, in which each equation directly represents an underlying geophysical process. The permafrost carbon feedback (PCF), melting of the Antarctic Ice Sheet (AIS) and Greenland Ice Sheet (GIS), and variability of the Indian Summer Monsoon (ISM) are modeled in this way. The other class uses survival analysis, whereby the underlying geophysical processes are represented more abstractly by a hazard event with an associated hazard rate. Slowdown of the Atlantic Meridional Overturning Circulation (AMOC) and Amazon rainforest dieback (AMAZ) are modeled in this second way. Both classes of model are probabilistic in META.

As a common measure of the extent to which different climate tipping elements have "tipped," we report an indicator of tipping point intensity on a scale from zero to one (supplementary materials). Zero corresponds to tipping elements whose state has been completely unaltered by climate change, and one corresponds to elements that have fully transitioned to an alternative state [as such, it is similar to the tipping state variable in (*48*)]. An indicator value of 0.5 means half-transformed in the case of process-based models (e.g., the ice sheets are half-melted), or a hazard event that is 50% likely in the case of the survival-analysis models. Although the damages from tipping points are ultimately most important, one advantage of this physical indicator is that it focuses attention on tipping processes, whose impacts are particularly poorly covered by existing economic studies. AMOC slowdown and AMAZ are leading examples (*40*). As can be seen in Fig. 5, methane action takes time to affect tipping point intensity. By 2050, however, the effect of methane action on tipping points becomes clearer. AIS and GIS tipping intensity decreases by 2.5%, the likelihood of AMOC slowdown decreases by 3%, the likelihood of AMAZ decreases by 8%, and ISM and PCF tipping intensity falls by 13%. After 2050, the effect of methane action on tipping point intensity



	Mean	Std. dev.	VaR 95%	VaR 99%	CVaR 95%	CVaR 99%
Tipping points						
Baseline	3.5	1.5	6.4	7.8	7.4	8.9
Methane action	3.1	1.4	5.9	7.3	6.7	8.0
No tipping points						
Baseline	3.1	1.4	5.9	7.3	6.8	8.1
Methane action	2.7	1.3	5.3	6.5	6.2	7.4

Fig. 4. Probability density functions for climate damages in 2050 with and without methane action. (**Top left**) Includes tipping points. (**Top right**) Omits tipping points. Specification includes market damages. The height of each bar corresponds to the frequency of observations in the corresponding bin. In both graphs, red indicates with methane action, and gray indicates without. (**Bottom**) Table reporting associated distribution statistics (percentage of consumption lost in 2050 because of climate change).

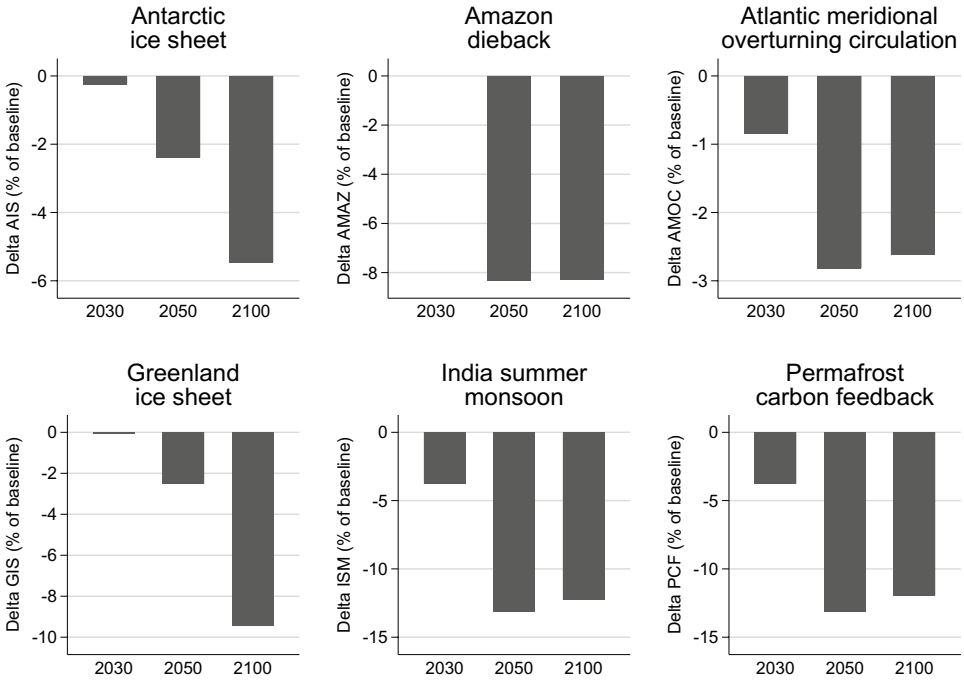


Fig. 5. Percentage change in tipping point intensity due to methane action for six tipping points. Background emissions scenario is SSP2-4.5. Tipping point indicators are described in the materials and methods.

generally mimics the effect of methane action on global mean temperature by plateauing. The exceptions are AIS and GIS tipping intensity, which are subject to considerable inertia. Therefore, methane action has its largest effect on these by the end of the century, reducing tipping intensity by ~5.5% (AIS) and 9.5% (GIS). Figure S8 in the supplementary materials again repeats the analyses in this section under the SSP3-7.0 background scenario, finding larger reductions in tipping point intensity than under SSP2-4.5.

SC-CH_4 and noncooperative action

The top panel of Fig. 6 shows probability density functions for the SC-CH_4 in 2020. These estimates reflect the marginal damage from 1 metric ton of CH_4 (tCH_4) emitted in 2020. The distribution has a strong positive skew with a mean of \$7381/tCH_4, a median of \$6837, and a 90% confidence interval of \$3824 to \$11,548. We also plot the corresponding probability density function excluding all tipping points, which has a mean of \$6307. Thus, the tipping points we include increase the mean SC-CH_4 by 17%. If the more speculative ocean methane hydrates tipping point is included, the SC-CH_4 in 2020 is \$23,319, which is 270% higher than the SC-CH_4 without tipping points. This is a far bigger impact of the ocean methane hydrates tipping point than on the SC-CO_2 (*40*). In the supplementary materials, we report the results of global sensitivity analysis of the SC-CH_4 with respect to climate, damage, and tipping point parameters.

Table S1 in the supplementary materials reports values for the SC-CH_4 in future years. These increase because of rising incomes and temperatures. For an emission in 2030, the mean SC-CH_4 in our preferred specification is \$9988/tCH_4, rising to \$13,824 for an emission in 2050 and \$27,044 for an emission in 2100. This corresponds to a

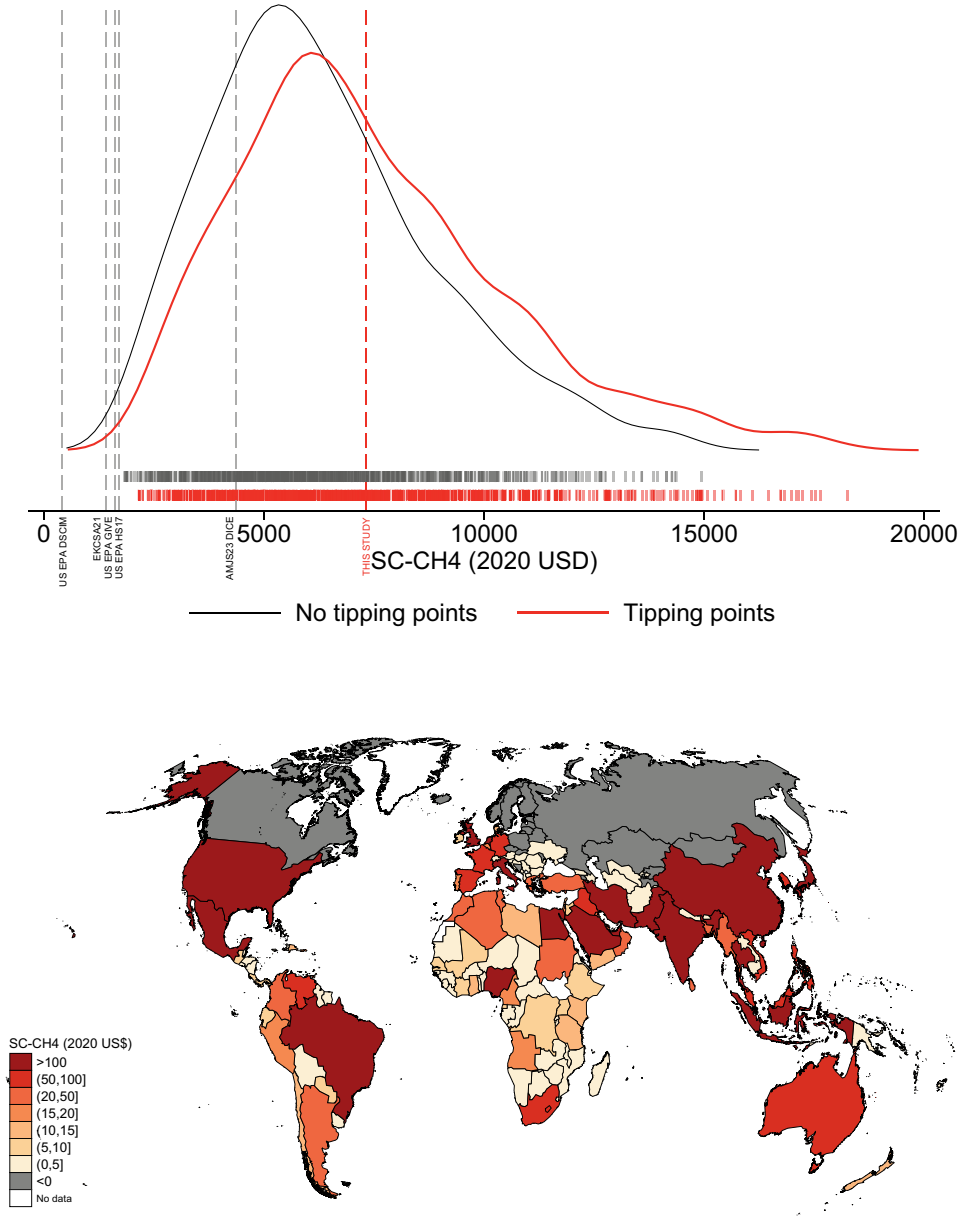


Fig. 6. Probability density function for the global SC-CH_4 in 2020 and map of expected values of national SC-CH_4 values. (**Top**) Probability density function. (**Bottom**) Map of expected values. Social costs are reported in 2020 US dollars for a 2020 emissions pulse in the SSP2-4.5 emissions scenario. Welfare changes are normalized to global mean consumption per capita for global SC-CH_4 and national mean consumption per capita for national SC-CH_4 values. Specification includes nonmarket damages, EMUC = 1.05, and PRTP = 0.5%. Tipping points are included in map. Previous estimates of the global SC-CH$_4$ are superimposed on the probability density function.

compound annual real growth rate of 1.6% between 2020 and 2100. The supplementary materials also reports values for the SC-CO_2. The mean SC-CO_2 in 2020, including tipping points, is \$309/tCO_2, rising to \$485 in 2050 and \$688 in 2100, a compound annual real growth rate of 1%. The ratio of the SC-CH_4 to the SC-CO_2 is 24 in 2020, rising to 39 in 2100. This ratio can be compared with the GWP of methane, which is 27 to 30 when measured over 100 years and 81 to 83 when measured over 20 years (*49*). Thus, the use of GWP100 would slightly overstate the value of methane abatement in the short run and slightly understate it in the long run, with the crossover period being about mid-century (SC-CH_4/SC-CO_2 = 28.5 in 2050). The use of GWP20 would greatly overstate the value of methane abatement. The discount rate is a factor. Building on the expert survey results of (*50*) and recent changes in practice (*51*), we set a pure rate of time preference (ρ) of 0.5% and an elasticity of marginal utility of consumption (η) of 1.05. With global economic growth on SSP2-4.5 of $g \sim 2\%$, this yields a discount rate of $r = \rho + \eta g \sim 2.5\%$ (because g varies over time, country, and Monte Carlo draw, this is approximate). Results from table S2 in the supplementary materials show that if we instead set ρ to 1% and η to 1.5, the ratio is 39 in 2020, higher than GWP100 but still much lower than GWP20. GWP100 appears to be a reasonable approximation of the relative social value of methane action, with two caveats: (i) it is a better approximation at low discount rates; (ii) these results are obtained from META and are hence limited to an analysis of avoided climate damages, rather than air-quality cobenefits.

The bottom panel of Fig. 6 reports the SC-CH_4 at the national level. Whereas we provide country-level estimates of the SC-CH_4, previous studies such as (*52*) have provided country-level estimates of the SC-CO_2. Formally, the SC-GHG is the change in social welfare from emitting 1 metric ton of the GHG, divided by the marginal utility of consumption to convert it into money units. The choice of consumption level for the denominator is important. The higher this is, the smaller is the marginal utility of consumption under the standard assumption of diminishing marginal utility, the smaller is the denominator in turn, and the higher is the SC-GHG in money terms. The use of global mean consumption per capita normalizes national welfare changes to the same level and effectively disaggregates the global SC-GHG that each country would refer to under full international cooperation into national shares. If instead each country's welfare change is normalized on the basis of its own national consumption per capita, this yields an estimate of climate damages to that country valued at its own income level. This corresponds to the SC-GHG that each country would use under noncooperation. Because we already analyzed the implications of cooperative global methane abatement for global inequality above, we take the latter approach. Studying noncooperative incentives is also highly relevant in current geopolitical conditions.

Negative SC-CH_4 values are reported for 20 primarily high-latitude countries, consistent with previous results generated from nonlinear temperature damage functions (*53*), even allowing for additional damage channels in META. These countries would not have a private incentive to reduce methane emissions at the margin today, assuming that META's damage estimates are not biased toward being too small. One hundred fifty-nine countries have a positive SC-CH_4, and 17 of these have a national SC-CH_4 above \$100/tCH_4. The three highest national values are for China (\$536/tCH_4), India (\$548/tCH_4), and the

US ($1112/tCH$_4$). The size of the US SC-CH$_4$ is notable because it implies that substantial methane abatement would be optimal even if the US returned to the first Trump administration practice of valuing only domestic damages from domestic methane abatement, as seems highly likely. The European Union (EU27) SC-CH$_4$, computed as the sum of its member states' national SC-CH$_4$ values, is $631/tCH$_4$, greater than the values of China and India.

Discussion

The benefits of global methane action look so much larger than the costs that the economic case for action is clear. Implementation of the Global Methane Pledge might come at a negative cost, according to leading estimates from (*43*) and (*10*). Even if one is skeptical of negative abatement-cost estimates, our modeling suggests much higher benefits from avoided climate damages and improved air quality. An important part of the global benefit of methane abatement is that avoided climate damages are higher in low- and middle-income countries, making the climate benefits of methane abatement pro-poor. Another important factor is that methane abatement reduces climate risk to the economy, something the risk-averse social planner values. Implementation of the Global Methane Pledge supposes a high degree of cooperation between countries, but in a world of fracturing geopolitics, how much methane abatement should countries undertake in their own self-interest? We analyzed this by estimating each country's domestic marginal damages and found that methane abatement at marginal costs of several hundred US dollars per tCH$_4$ is economical for several major players such as China, the EU, India, and the US (where the domestic SC-CH$_4$ is as high as $1112/tCH$_4$). Lastly, although the best basis for comparing methane and CO$_2$ abatement should be relative social costs, we have found that the convention of using GWP100 is a reasonable approximation.

How do our estimates of climate benefits from methane abatement compare with previous literature? Estimates of climate benefits are famously sensitive to the discount rate (albeit less so for methane than CO$_2$ because of its shorter atmospheric residence time), so any comparison needs to control for the discount rate as far as possible. It also needs to adjust for inflation, so we inflate values reported in previous studies where necessary using the US GDP deflator. In terms of total benefits of global methane action scenarios, the only point of comparison is (*23*). Using an aggregate damage function from a meta-analysis (*54*), they estimate total benefits of ~$3.5 trillion to ~$10 trillion for discount rates of 2 and 3%. Our estimates overlap this range significantly but are at the high end of it.

In terms of marginal damages, our global estimate of the 2020 SC-CH$_4$ (e.g., mean of $7381/tCH$_4$) can be compared with a wider range of studies. (*32*) calculate the SC-CH$_4$ by coupling a set of four Simple Climate Models to damages from the DICE and FUND IAMs. For a discount rate of 2.5% and on a high-emissions scenario (RCP8.5), they report a multimodel mean estimate of $1414/tCH$_4$ and a 90% confidence interval of $678 to $2454. But some model combinations yield values close to $4000—notably, using DICE damages instead of FUND. The US EPA (*34*) reports values of $470/tCH$_4$ with damages from the DSCIM IAM, $1600 with damages from the GIVE IAM, and $1700 with damages from (*54*), for a near-term discount rate of 2.5%. Using an updated version of DICE (including FaIR v2.0.0), (*35*) estimate $4360/tCH$_4$ within a range of $959 to $8829. Their central estimate is generated along the optimal emissions path by using a comparable discount rate to ours and a damage function calibrated on (*54*), such as the US EPA meta-analysis variant [but choosing a different model variant in (*54*) with higher damages for given temperature]. Their SC-CH$_4$ exhibits strong sensitivity to the emissions path, however, and when calculated on a business-as-usual (BAU) path, it is at the high end of their reported range, ~20% above our central estimate.

Overall, our estimated SC-CH$_4$ is higher than that in most of the existing literature, including estimates provided in the latest US EPA guidance. In particular, it extends the uncertainty range well above $10,000. With the discount rate controlled for and META using an almost identical climate model to (*34*) and (*35*) (FaIR), remaining differences in the SC-CH$_4$ most likely arise from different socioeconomic and emissions scenarios, different spatial disaggregation, and different modeling of damages (such as inclusion of tipping points in META). Our mean estimate looks consistent with those of (*35*) when one considers that we calculate the SC-CH$_4$ on an emissions path that is between their optimum and BAU. But the analysis in (*35*) is deterministic; we add quantification of uncertainty, which naturally spreads the distribution of the SC-CH$_4$ to higher values. Furthermore, META combines spatial disaggregation to the country level with welfare valuation, using diminishing marginal utility. Thus, SC-GHG estimates from META include so-called "distributional/equity weighting," unlike (*34*), or (*23*) for total benefits. The literature on the SC-CO$_2$ shows that estimates can increase strongly after distributional weighting (*55*, *56*) because greater value is placed on low-income countries' damages, which are relatively higher. This is not to be confused with the opposite role of diminishing marginal utility in normalizing welfare changes discussed above when computing national SC-CH$_4$ values.

This study has several limitations. Similar to any IAM, META is affected by multiple uncertainties ranging from what emissions or socioeconomic scenario to choose, through how the climate system responds to emissions, to, in particular, the damages climate change causes and the welfare value of those damages. We have sought to quantify those uncertainties through both scenario analysis and probabilistic modeling, but the resulting climate benefits could be misestimated and uncertainty is likely underestimated, on balance. On the other side of the ledger, the abatement costs provided for (*7*) that we use could also be misestimated. Our analysis only quantifies the impacts of methane emissions via climate damages and air pollution on health. But methane emissions have other economic impacts, for example, via ground-level ozone reducing agricultural productivity. Moreover, as is standard in economic benefit-cost analysis, we abstract from political costs. Still, the economic case for methane action is clear.

Materials and methods summary

We create a set of emissions scenarios that simulate methane action at the global level. These are anchored on successful achievement in 2030 of the 30% emissions reduction target in the Global Methane Pledge, followed by continued methane abatement post-2030. They are built with raw scenario data from the IAMC 1.5°C Scenario Explorer (*57*). These methane action scenarios are then superposed on a set of background emissions scenarios, which (i) set baseline methane emissions for comparison and (ii) set emissions of CO$_2$ and the other GHGs and radiative forcing agents, which are held constant, whereas methane emissions are varied. We consider a range of background emissions and forcing scenarios: a low-emissions scenario that limits the increase in global mean temperature to 1.5°C above preindustrial with low or no overshoot, a "middle-of-the-road" emissions and socioeconomic scenario (SSP2-4.5), and a higher-emissions scenario characterized by "regional rivalry" in economic and/or political relations (SSP3-7.0). This approach assumes that methane action can be independent of action on other GHGs to a first-order approximation, which is supported by evidence that targeted methane abatement measures, which would be sufficient to achieve the Global Methane Pledge, do not also reduce CO$_2$ (e.g., improved leak detection and repair in oil and gas infrastructure; landfill gas capture) (*7*). The opposite is unlikely to be true: CO$_2$ abatement would also reduce methane (e.g., switching to renewables reduces fossil fuel use, which in turn reduces methane emissions from upstream fossil fuel production and distribution).

The second step is to take the scenarios and run them through the META IAM to project climate and economic effects. For this paper, we implemented several improvements and updates to the META model, including converting the model to the open-source Mimi.jl platform; integrating the FaIRv2.0.0 climate model (*42*); reestimating our downscaling functions from global to national temperature on the CMIP6 dataset

instead of CMIP5; replacing the original West Antarctic Ice Sheet tipping point module with an AIS tipping point module (*41*), which replicates the impulse responses of major ice sheet models (*58*); revising hazard rates for the AMOC and Ocean Methane Hydrates tipping point modules; revising estimates of the country-level temperature damage function parameters, still using (*53*); and improving the representation of nonmarket damages using more up-to-date literature (*54*). These changes, as well as the full set of equations making up the META model, are described in the supplementary materials.

The third step brings in additional, scenario-consistent data on methane abatement costs and air-quality cobenefits. For the former, we use sectoral estimates of methane abatement costs from (*43*) and (*10*), which were used in the Global Methane Assessment (*7*), to construct streams of undiscounted abatement costs from 2020 to 2050 that are compatible with the quantity of abatement required to shift methane emissions from current trends to the central methane action scenario. We also use the methane MAC curve estimated by (*8*) to construct a further stream of undiscounted abatement costs over the same period. For the latter, we again use data from the Global Methane Assessment (*7*), from which we obtain estimates of premature deaths from respiratory and cardiovascular diseases due to tropospheric ozone generated by methane emissions. We convert these to marginal economic damages using the VSL and assumptions about the income elasticity of a VSL. We then multiply this quantity by the emissions reductions from our methane action scenarios, giving us a stream of undiscounted air-quality cobenefits.

REFERENCES AND NOTES

1. Intergovernmental Panel on Climate Change (IPCC), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, V. Masson-Delmotte *et al.*, Eds. (Cambridge Univ. Press, 2021).
2. D. Shindell *et al.*, The methane imperative. *Front. Sci.* **2**, 1349770 (2024). doi: 10.3389/fsci.2024.1349770
3. J. G. Canadell *et al.*, "Global Carbon and other Biogeochemical Cycles and Feedbacks" in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, V. Masson-Delmotte *et al.*, Eds. (Cambridge Univ. Press, 2021), pp. 673–816.
4. M. Saunois *et al.*, Global methane budget 2000–2020. *Earth Syst. Sci. Data* **17**, 1873–1958 (2025). doi: 10.5194/essd-17-1873-2025
5. L. Feng, P. I. Palmer, S. Zhu, R. J. Parker, Y. Liu, Tropical methane emissions explain large fraction of recent changes in global atmospheric methane growth rate. *Nat. Commun.* **13**, 1378 (2022). doi: 10.1038/s41467-022-28989-z; pmid: 35297408
6. Z. Qu *et al.*, Inverse modeling of 2010–2022 satellite observations shows that inundation of the wet tropics drove the 2020–2022 methane surge. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2402730121 (2024). doi: 10.1073/pnas.2402730121; pmid: 39316054
7. United Nations Environment Programme (UNEP), Climate and Clean Air Coalition, "Global Methane Assessment: Benefits and Costs of Mitigating Methane Emissions" (UNEP, 2021).
8. M. Harmsen *et al.*, Long-term marginal abatement cost curves of non-CO_2 greenhouse gases. *Environ. Sci. Policy* **99**, 136–149 (2019). doi: 10.1016/j.envsci.2019.05.013
9. M. Harmsen *et al.*, The role of methane in future climate strategies: Mitigation potentials and climate impacts. *Clim. Change* **163**, 1409–1425 (2020). doi: 10.1007/s10584-019-02437-2
10. L. Höglund-Isaksson, A. Gómez-Sanabria, Z. Klimont, P. Rafaj, W. Schöpp, Technical potentials and costs for reducing global anthropogenic methane emissions in the 2050 timeframe – Results from the gains model. *Environ. Res. Commun.* **2**, 025004 (2020). doi: 10.1088/2515-7620/ab7457
11. I. B. Ocko *et al.*, Acting rapidly to deploy readily available methane mitigation measures by sector can immediately slow global warming. *Environ. Res. Lett.* **16**, 054042 (2021). doi: 10.1088/1748-9326/abf9c8
12. I. B. Ocko *et al.*, Unmask temporal trade-offs in climate policy debates. *Science* **356**, 492–493 (2017). doi: 10.1126/science.aaj2350; pmid: 28473552
13. C. S. Malley *et al.*, A roadmap to achieve the global methane pledge. *Environ. Res. Climate* **2**, 011003 (2023). doi: 10.1088/2752-5295/acb4b4
14. R. T. Pierrehumbert, Short-lived climate pollution. *Annu. Rev. Earth Planet. Sci.* **42**, 341–379 (2014). doi: 10.1146/annurev-earth-060313-054843
15. European Commission, Global Methane Pledge (2021); https://globalmethanepledge.ccacoalition.org/.
16. J. Rogelj, M. Meinshausen, M. Schaeffer, R. Knutti, K. Riahi, Impact of short-lived non-CO_2 mitigation on carbon budgets for stabilizing global warming. *Environ. Res. Lett.* **10**, 075001 (2015). doi: 10.1088/1748-9326/10/7/075001
17. W. J. Collins *et al.*, Increased importance of methane reduction for a 1.5 degree target. *Environ. Res. Lett.* **13**, 054003 (2018). doi: 10.1088/1748-9326/aab89c
18. D. Shindell *et al.*, Simultaneously mitigating near-term climate change and improving human health and food security. *Science* **335**, 183–189 (2012). doi: 10.1126/science.1210026; pmid: 22246768
19. J. K. Shoemaker, D. P. Schrag, M. J. Molina, V. Ramanathan, Climate change. What role for short-lived climate pollutants in mitigation policy? *Science* **342**, 1323–1324 (2013). doi: 10.1126/science.1240162; pmid: 24337280
20. A. Stohl *et al.*, Evaluating the climate and air quality impacts of short-lived pollutants. *Atmos. Chem. Phys.* **15**, 10529–10566 (2015). doi: 10.5194/acp-15-10529-2015
21. M. T. Lund *et al.*, A continued role of short-lived climate forcers under the shared socioeconomic pathways. *Earth Syst. Dyn.* **11**, 977–993 (2020). doi: 10.5194/esd-11-977-2020
22. S. J. Smith *et al.*, Impact of methane and black carbon mitigation on forcing and temperature: A multi-model scenario analysis. *Clim. Change* **163**, 1427–1442 (2020). doi: 10.1007/s10584-020-02794-3
23. B. Cael, P. Goodwin, Global methane pledge versus carbon dioxide emission reduction. *Environ. Res. Lett.* **18**, 104015 (2023). doi: 10.1088/1748-9326/acf8dd
24. A. Hu, Y. Xu, C. Tebaldi, W. M. Washington, V. Ramanathan, Mitigation of short-lived climate pollutants slows sea-level rise. *Nat. Clim. Chang.* **3**, 730–734 (2013). doi: 10.1038/nclimate1869
25. S. Fankhauser, The social costs of greenhouse gas emissions: An expected value approach. *Energy J.* **15**, 157–184 (1994). doi: 10.5547/ISSN0195-6574-EJ-Vol15-No2-9
26. M. Hoel, I. Isaksen, in *Control and Game-Theoretic Models of the Environment*, C. Carraro, J. A. Filar, Eds. (Birkhäuser, 1995), pp. 89–105.
27. R. S. Tol, The marginal costs of greenhouse gas emissions. *Energy J.* **20**, 61–81 (1999). doi: 10.5547/ISSN0195-6574-EJ-Vol20-No1-4
28. C. Hope, The marginal impact of CO_2 from PAGE2002: An integrated assessment model incorporating the IPCC's five reasons for concern. *Integr. Assess. J.* **6**, 19–56 (2006).
29. A. L. Marten, S. C. Newbold, Estimating the social cost of non-CO_2 GHG emissions: Methane and nitrous oxide. *Energy Policy* **51**, 957–972 (2012). doi: 10.1016/j.enpol.2012.09.073
30. D. T. Shindell, J. S. Fuglestvedt, W. J. Collins, The social cost of methane: Theory and applications. *Faraday Discuss.* **200**, 429–451 (2017). doi: 10.1039/C7FD00009J; pmid: 28581559
31. M. C. Sarofim, S. T. Waldhoff, S. C. Anenberg, Valuing the ozone-related health benefits of methane emission controls. *Environ. Resour. Econ.* **66**, 45–63 (2017). doi: 10.1007/s10640-015-9937-6
32. F. C. Errickson, K. Keller, W. D. Collins, V. Srikrishnan, D. Anthoff, Equity is more important for the social cost of methane than climate uncertainty. *Nature* **592**, 564–570 (2021). doi: 10.1038/s41586-021-03386-6; pmid: 33883735
33. J. Sampedro, S. Waldhoff, M. Sarofim, R. Van Dingenen, Marginal damage of methane emissions: Ozone impacts on agriculture. *Environ. Resour. Econ.* **84**, 1095–1126 (2023). doi: 10.1007/s10640-022-00750-6; pmid: 39376640
34. US Environmental Protection Agency (EPA), "Report on the Social Cost of Greenhouse Gases: Estimates Incorporating Recent Scientific Advances" (2023).
35. C. Azar, J. G. Martín, D. J. Johansson, T. Sterner, The social cost of methane. *Clim. Change* **176**, 71 (2023). doi: 10.1007/s10584-023-03540-1
36. R. Schmalensee, Comparing greenhouse gases for policy purposes. *Energy J.* **14**, 245–255 (1993). doi: 10.5547/ISSN0195-6574-EJ-Vol14-No1-10
37. D. S. Mallapragada, B. K. Mignone, A theoretical basis for the equivalence between physical and economic climate metrics and implications for the choice of global warming potential time horizon. *Clim. Change* **158**, 107–124 (2020). doi: 10.1007/s10584-019-02486-7
38. N. Stern, *The Economics of Climate Change: The Stern Review* (Cambridge Univ. Press, 2007).
39. S. Dietz, C. Hepburn, Benefit–cost analysis of non-marginal climate and energy projects. *Energy Econ.* **40**, 61–71 (2013). doi: 10.1016/j.eneco.2013.05.023
40. S. Dietz, J. Rising, T. Stoerk, G. Wagner, Economic impacts of tipping points in the climate system. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2103081118 (2021). doi: 10.1073/pnas.2103081118; pmid: 34400500
41. S. Dietz, F. Koninx, Economic impacts of melting of the Antarctic Ice Sheet. *Nat. Commun.* **13**, 5819 (2022). doi: 10.1038/s41467-022-33406-6; pmid: 36192387
42. N. J. Leach *et al.*, Fairv2.0.0: A generalized impulse response model for climate uncertainty and future scenario exploration. *Geosci. Model Dev.* **14**, 3007–3036 (2021). doi: 10.5194/gmd-14-3007-2021
43. EPA, "Global Non-CO_2 Greenhouse Gas Emission Projections & Marginal Abatement Cost Analysis: Methodology Documentation" (Technical Report, EPA-430-R-19-012, 2019).
44. UNEP, "Executive summary, Emissions Gap Report 2023: Broken Record – Temperatures hit new highs, yet world fails to cut emissions (again)" (2023).

45. X. Lan, K. W. Thoning, E. J. Dlugokencky, "Trends in globally-averaged CH_4, N_2O, and SF_6 determined from NOAA Global Monitoring Laboratory measurements" (Version 2024-02, 2022).
46. S. Dietz, in *Handbook of the Economics of Climate Change*, vol. 1, L. Barrage, S. Hsiang, Eds. (Elsevier, 2024), pp. 1–51.
47. M. J. Kotchen, J. A. Rising, G. Wagner, The costs of "costless" climate mitigation. *Science* **382**, 1001–1003 (2023). doi: 10.1126/science.adj2453; pmid: 38033066
48. Y. Cai, T. S. Lontzek, The Social Cost of Carbon with Economic and Climate Risks. *J. Polit. Econ.* **127**, 2684–2734 (2019) doi: 10.1086/701890
49. P. Forster *et al.*, "The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity" in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, V. Masson-Delmotte *et al.*, Eds. (Cambridge Univ. Press, 2021), pp. 923–1054.
50. M. Drupp, M. Freeman, B. Groom, F. Nesje, Discounting disentangled. *Am. Econ. J. Econ. Policy* **10**, 109–134 (2018). doi: 10.1257/pol.20160240
51. US Office of Management and Budget, "Circular No. A-4" (2023).
52. K. Ricke, L. Drouet, K. Caldeira, M. Tavoni, Country-level social cost of carbon. *Nat. Clim. Chang.* **8**, 895–900 (2018). doi: 10.1038/s41558-018-0282-y
53. M. Burke, S. M. Hsiang, E. Miguel, Global non-linear effect of temperature on economic production. *Nature* **527**, 235–239 (2015). doi: 10.1038/nature15725; pmid: 26503051
54. P. H. Howard, T. Sterner, Few and not so far between: A meta-analysis of climate damage estimates. *Environ. Resour. Econ.* **68**, 197–225 (2017). doi: 10.1007/s10640-017-0166-z
55. F. Dennig, M. B. Budolfson, M. Fleurbaey, A. Siebert, R. H. Socolow, Inequality, climate impacts on the future poor, and carbon prices. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15827–15832 (2015). doi: 10.1073/pnas.1513967112; pmid: 26644560
56. D. Anthoff, J. Emmerling, Inequality and the social cost of carbon. *J. Assoc. Environ. Resour. Econ.* **6**, 243–273 (2019). doi: 10.1086/701900
57. D. Huppmann *et al.*, "IAMC 1.5°C Scenario Explorer and Data hosted by IIASA" (Integrated Assessment Modeling Consortium & International Institute for Applied Systems Analysis, 2019).
58. A. Levermann *et al.*, Projecting Antarctica's contribution to future sea level rise from basal ice shelf melt using linear response functions of 16 ice sheet models (LARMIP-2). *Earth Syst. Dyn.* **11**, 35–76 (2020). doi: 10.5194/esd-11-35-2020
59. T. Stoerk, J. Rising, D. Shindell, S. Dietz, Replication package for "Global methane action pays for itself at least six times over" (v1.0), Zenodo (2025). doi: 10.5281/zenodo.15363212

ACKNOWLEDGMENTS

We thank T. Carleton for detailed discussant's comments; Z. Nicholls for sharing emissions data; and audiences at the Asian Development Bank, Barcelona School of Economics, Environmental Defense Fund, European Commission, IE Business School, LSE, NOVA Business School, Oxford, Resources for the Future, ASSA, AERE, AERNA, Atlantic Workshop on Energy and Environmental Economics, CEPR Paris Symposium, and EAERE for their helpful feedback. This work does not necessarily represent the views of the National Bank of Belgium or the Eurosystem. **Funding:** D.S. acknowledges the financial support of the Global Methane Hub. S.D. acknowledges the financial support of the Grantham Foundation for the Protection of the Environment. **Author contributions:** S.D., J.R., and T.S. conceived the study. D.S. and T.S. constructed the scenarios with input from J.R. and S.D. S.D., J.R., and T.S. adapted the model, with J.R. leading on software development and implementation. S.D., J.R., and T.S. produced the results and analyzed the data. S.D., J.R., D.S., and T.S. wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data underlying the Model for Economic Tipping point Analysis (META) used in this study are fully documented in the supplementary materials and draw on public sources. The version of META used in this study is freely available on GitHub (https://github.com/openmodels/META/releases/tag/META-2025). All climate data and code necessary to reproduce the analytical steps and results in this work are available from Zenodo (*59*).

License information: Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adu7392
Materials and Methods; Supplementary Text; META Model Description; Figs. S1 to S17; Tables S1 to S8; References (*60–118*)

Submitted 19 November 2024; accepted 21 August 2025

10.1126/science.adu7392

Science
JOURNALS AAAS



Publish your research in the *Science* family of journals

The *Science* family of journals (*Science, Science Advances, Science Immunology, Science Robotics, Science Signaling*, and *Science Translational Medicine)* are among the most highly-regarded journals in the world for quality and selectivity. Our peer-reviewed journals are committed to publishing cutting-edge research, incisive scientific commentary, and insights on what's important to the scientific world at the highest standards.

Submit your research today!

Learn more at **Science.org/journals**

RESEARCH ARTICLE SUMMARY

NEUROIMMUNOLOGY

Regional encoding of enteric nervous system responses to microbiota and type 2 inflammation

Peng Tan†, Alok Jaiswal†, Shane P. Murphy, Eric M. Brown, Hailey Wheeler, Chien-Wen Su, Emily P. Finan, Guadalupe J. Jasso, Hai Ning Shi, Daniel B. Graham, Toni M. Delorey, Jacques Deguine*, Ramnik J. Xavier*

Full article and list of author affiliations: https://doi.org/10.1126/science.adr3545

INTRODUCTION: The enteric nervous system (ENS) is a branch of the autonomous nervous system composed of neurons located across two plexi of the intestinal tract. The ENS can function semiautonomously to regulate digestion and peristalsis, through a circuit involving five main neuronal types: (i) sensory neurons activate (ii) excitatory and (iii) inhibitory motor neurons through (iv) ascending and descending interneurons, which concomitantly activate (v) secretomotor and vasodilator neurons to regulate epithelial secretion and absorption. Accumulating evidence suggests that, in addition to working through this core circuit, the ENS plays a key role in responses to bacteria, allergens, and parasites through extensive cross-talk with the immune system.

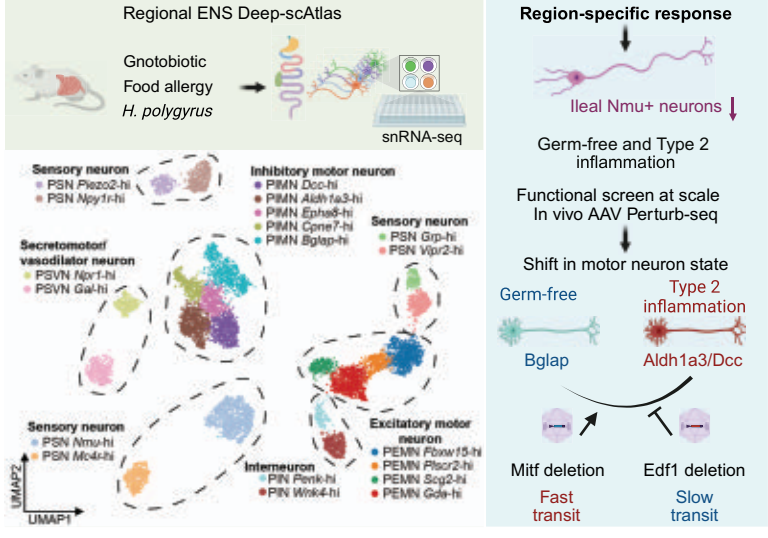
RATIONALE: The gastrointestinal tract is constantly exposed to environmental changes driven by the microbiota, pathogens, and immunity, yet our understanding of the ENS's adaptation to these fluctuations remains limited, in part because of difficulties in isolating and profiling enteric neurons. To address this gap, we generated gnotobiotic mouse reporters and colonized them with distinct microbiomes, and in parallel, studied the ENS of animals exposed to type 2 inflammation, including food allergy and *Heligmosomoides polygyrus* infection. In each of these settings, we deeply profiled the ENS across distinct segments of the intestine—duodenum, jejunum, ileum, and colon—to define its response to perturbation.

RESULTS: Plate-based single-nucleus RNA profiling of the ENS captured a total of 7640 enteric neurons with more than 6000 median genes per cell. This atlas highlights two main axes of diversity across enteric neurons. First, we observed a functional specification of sensory neurons across distinct cell types characterized by distinct neuropeptides and receptors, including neuromedin U (Nmu)–expressing neurons that respond to type 2 cytokines and leukotrienes. Second, we observed transcriptomic gradients across both subsets of motor neurons (inhibitory and excitatory), likely encompassing distinct cell states that can be modulated by environmental challenges. Notably, these changes were highly coordinated across distinct perturbations, with changes in the frequency of Nmu^+ neurons and a shift of inhibitory motor neurons between *Bglap* and *Dcc*/*Aldh1a3* populations occurring in animals that were allergic, parasite-infected, and born germ-free.

To define the programs that contribute to neuronal adaptation, we next used adeno-associated virus (AAV)–PHP.S viral vectors—which efficiently target all peripheral neurons—to delete genes of interest in *Cas9*-expressing animals. We first delivered pools of guides and performed single-nucleus RNA sequencing of infected neurons to identify genes associated with changes in neuronal cell states. Lastly, we deleted candidate regulators in individual animals and validated two genes, *Edf1* and *Mitf*, as regulators of both inhibitory motor neuron cell state and total gastrointestinal transit time. More specifically, deletion of *Edf1* resulted in an expansion of $BGLAP^+$ neurons among infected cells and increased transit time, whereas deletion of *Mitf* expanded DCC^+ neurons and shortened transit time.

CONCLUSION: Our studies provide a detailed transcriptomic atlas of the ENS during distinct environmental challenges and reveal a set of coordinated responses to perturbations. This work directly links ENS cell states to changes in intestinal physiology and provides a blueprint for future functional studies of the ENS in health and disease. □

*Corresponding authors: jdeguine@broadinstitute.org (J.D.); xavier@molbio.mgh.harvard.edu (R.J.X.) †These authors contributed equally to this work. **Cite this article as P. Tan *et al.*, *Science* 390, eadr3545 (2025). DOI: 10.1126/science.adr3545**



Enteric neurons across sites and perturbations. Enteric neuron single-nucleus RNA sequencing (snRNA-seq) was performed on the indicated segments and models, resulting in a single-cell (sc) atlas of 7640 deeply profiled neurons (bottom left). *Nmu*-hi sensory neurons and inhibitory motor neuron states were highly responsive to perturbations (top right), and AAV-based gene deletions demonstrated a role for *Mitf* and *Edf1* in modulating inhibitory motor neurons and gastrointestinal transit time (bottom right). PSN, putative sensory neurons; PEMN, putative excitatory motor neurons; PIMN, putative inhibitory motor neurons; PIN, putative interneurons; PSVN, putative secretomotor/vasodilator neurons.

NEUROIMMUNOLOGY

Regional encoding of enteric nervous system responses to microbiota and type 2 inflammation

Peng Tan[1,2,3]†, Alok Jaiswal[2]†, Shane P. Murphy[1], Eric M. Brown[2,4], Hailey Wheeler[2], Chien-Wen Su[5], Emily P. Finan[1], Guadalupe J. Jasso[2], Hai Ning Shi[5,6], Daniel B. Graham[2,3,4,6], Toni M. Delorey[1], Jacques Deguine[1,2]*, Ramnik J. Xavier[1,2,7,4,8,6]*

Enteric neurons are essential regulators of intestinal physiology, yet their responses to varying microbial and immune environments along the intestinal tract and or during challenges remain poorly understood. In this study, we regionally profiled enteric neurons across gnotobiotic, allergic, and parasite-infected mice. Timing and complexity of microbial perturbations and type 2 inflammation result in motor neuron state shifts and alter multiple functionally distinct sensory neurons, including interleukin-13– and leukotriene-responsive *Nmu*-hi cells and *Grp*-hi neurons, which expand in germ-free colonic tissue and interact with *Grpr*$^+$ interstitial cells of Cajal. Leveraging adeno-associated virus–based Perturb-seq, we identified *Edf1* and *Mitf* as controllers of motor neuron state transition and gastrointestinal transit time, directly linking enteric neuron states to physiology.

The intestine is densely innervated by both extrinsic neurons and intrinsic neuronal populations that comprise the enteric nervous system (ENS) and act as semiautonomous circuits across the submucosal (Meissner's) plexus and the myenteric (Auerbach's) plexus (*1*–*5*). Intestinal peristalsis is initiated by intrinsic primary afferent neurons (IPANs) acting as sensory neurons, which activate excitatory and inhibitory motor neurons through ascending and descending interneurons, respectively. The ENS also contains secretomotor-vasodilator neurons (*6*). ENS dysregulation has been implicated in various intestinal disorders, most notably Hirschsprung's disease, in which *rearranged during transfection* (*RET*) mutations lead to defects in ENS development and bowel obstruction (*7*).

In addition to the peristaltic reflex, the ENS senses multiple environmental factors to regulate intestinal physiology. Indeed, sensing of microbiota by the ENS regulates intestinal motility and blood glucose (*8*, *9*), and ENS signals can promote gut healing (*10*) and antimicrobial peptide secretion (*11*). The ENS responds to psychological states (*12*), and subset-specific activation can modulate intestinal transit and feeding (*13*, *14*). During infection, the enteric neuron–derived peptide neuromedin U (Nmu) promotes parasite expulsion and tissue protection through binding to Nmur1 on type 2 innate lymphoid cells (ILC2s) (*15*, *16*) and eosinophils (*17*).

Although advances in single-nucleus RNA sequencing (snRNA-seq) have facilitated the characterization of the ENS (*18*), our understanding of its responses to challenges remains limited. Here, we deeply sequenced 7640 enteric neurons across intestinal regions and perturbations (Fig. 1A). We uncovered dynamics in both neuronal subsets and transcriptional states that are region- and perturbation-specific, including changes in sensory neuron abundance and shifts in the transcriptional states of excitatory and inhibitory motor neurons. We next designed and used an adeno-associated virus (AAV)–based in vivo Perturb-seq (*19*) approach to evaluate a pool of genes and individually validate *Edf1* and *Mitf* as regulators of inhibitory motor neuron programs and, concomitantly, total gastrointestinal transit time. Our findings define how microbiome and inflammation affect the ENS and, downstream, alter intestinal physiology.

Characterization of ENS–cell type diversity in multiple perturbations

To profile the ENS, we utilized the neuron-specific *Uchl1* (also known as PGP9.5) reporter mouse *Uchl1-H2B-mCherry/GFPgpi* (*20*). In these animals, neurons express mCherry fused to histone H2B, enabling sorting of the nuclei of enteric neurons. We observed that *Uchl1*-reporter$^+$ cells were spatially positioned across both submucosal and myenteric plexuses and validated the reporter's specificity against the known neuronal markers MAP2 (*21*) (fig. S1A) and HuC/D (fig. S1B).

Given the pivotal role of the microbiome in ENS development and gut physiology, we sought to establish mouse models to investigate responses to a gradient of microbial diversity (Fig. 1A and fig. S1C). We first rederived *Uchl1* mice to germ-free (GF) conditions and analyzed GF animals as well as animals reconstituted with a single bacterium or a complex microbial community. As expected, a single-dose colonization of 10^8 colony-forming units (CFU) *Bacteroides thetaiotaomicron* (BT) (*22*) or altered Schaedler flora (ASF; a defined community of eight bacterial species) (*23*) for 4 weeks effectively restored the number of immunoglobulin A (IgA$^+$) plasma cells (*24*) to amounts akin to those observed in sex- and age-matched specific pathogen–free (SPF) mice (fig. S1D).

We also examined responses to type 2 inflammation (Fig. 1A and fig. S1, E and I). For food allergy, mice were sensitized with ovalbumin (OVA) in alum administered intraperitoneally and subsequently challenged intragastrically. Challenge after sensitization resulted in increased expression of mast cell degranulation–associated chymase (*Mcpt1* and *Mcpt4*), tryptase (*Mcpt6*), carboxypeptidase (*Cpa3*), and T helper 2 (Th2) cytokines (*Il4*, *Il5*, and *Il13*), as well as elevated expression of total and OVA-specific immunoglobulin E (IgE) (fig. S1, F to H) but not anaphylaxis or diarrhea, which do not occur on a C57BL/6 background when challenged orally (*25*). To examine responses to a parasite, we studied animals 14 days postinfection with the exclusively enteric parasitic helminth *Heligmosomoides polygyrus* (*H. polygyrus*). Infected mice displayed elevated amounts of type 2 cytokines compared with controls, and parasite burdens were not affected by the presence of the *Uchl1* transgene (fig. S1, J and K).

From the gnotobiotic models, we took samples 4 cm in length from the terminal ileum and mid-colon (middle third segment) because these regions are characterized by high microbial loads (*26*); from the type 2 immunity models, we sampled the duodenum-jejunum and terminal ileum because allergic responses and inflammation were primarily induced in the small intestine. We isolated nuclei (*18*) from whole tissues and sorted 4′,6-diamidino-2-phenylindole (DAPI$^+$) Uchl1-H2B-mCherry$^+$ nuclei (fig. S2A) directly in lysis buffer for plate-based full-length SMART-seq2 (*27*), yielding a high-depth atlas of enteric neurons with >6000 median genes per cell (fig. S2B). We also integrated our dataset with our previously published study (table S1) (*18*). Dimensionality reduction and clustering of transcriptional profiles revealed five broad neuronal cell types (Fig. 1B and fig. S2C) defined as

[1]Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [2]Broad Institute of MIT and Harvard, Immunology Program, Cambridge, MA, USA. [3]Center for Drug Discovery and Translational Medicine, Hainan Academy of Medical Sciences, Hainan Medical University, Haikou, People's Republic of China. [4]Center for Computational and Integrative Biology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [5]Mucosal Immunology and Biology Research Center, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA, USA. [6]Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital, Boston, MA, USA. [7]The Gene Lay Institute of Immunology and Inflammation, Brigham and Women's Hospital, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. [8]Department of Molecular Biology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. ***Corresponding authors: jdeguine@broadinstitute.org (J.D.); xavier@molbio.mgh.harvard.edu (R.J.X.)** †These authors contributed equally to this work.

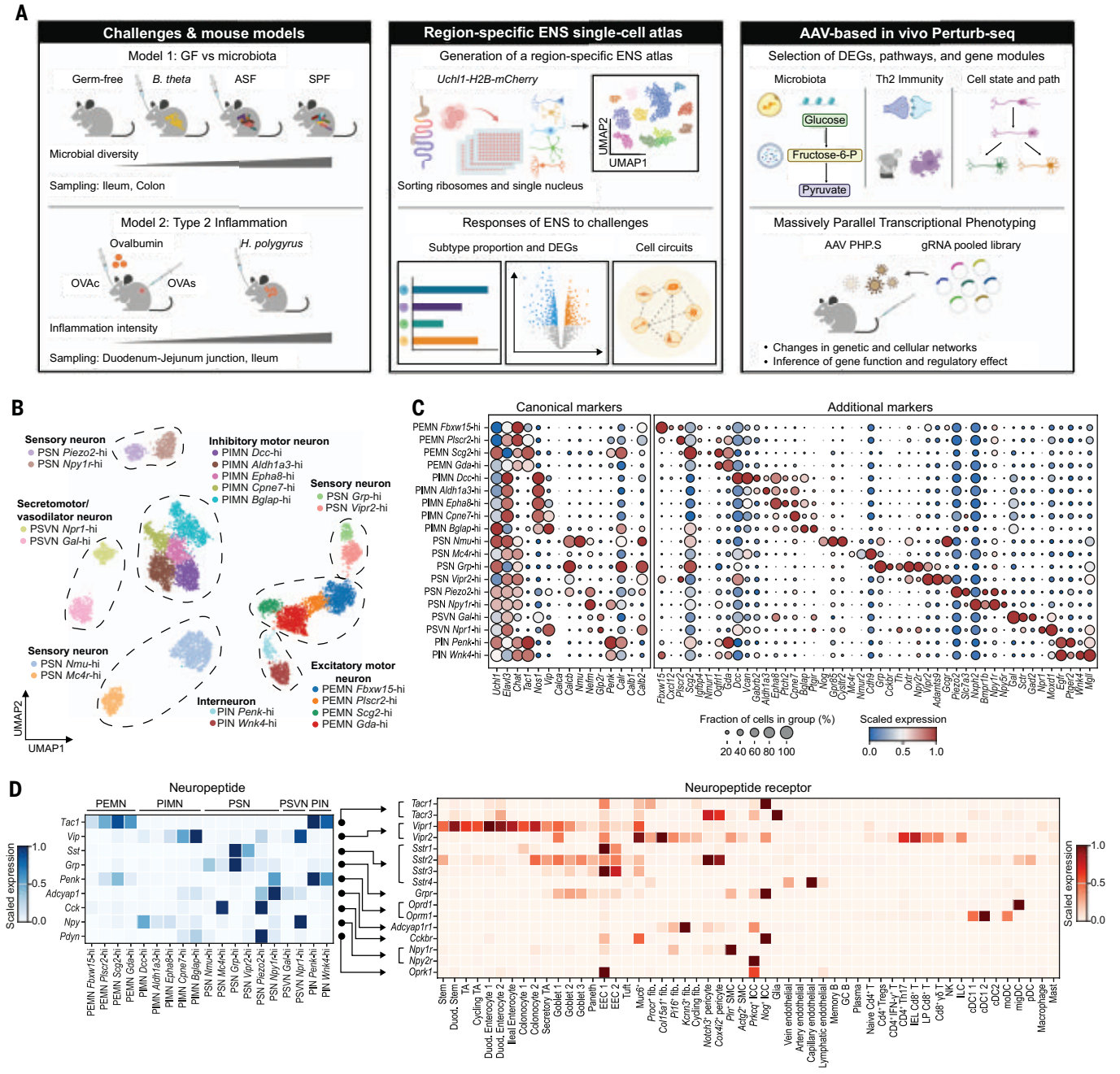


Fig. 1. Characterization of ENS–cell type diversity in multiple perturbations. (**A**) Overview of the study design. A neuron-specific *Uchl1-H2B-mCherry* reporter mouse was used to investigate ENS responses to commensal microbes (model 1) and type 2 inflammation (model 2), induced by either OVA-induced food allergy (mild) or *H. polygyrus* infection (strong). A region-specific ENS single-cell atlas was generated for the duodenum-jejunum (DJ) junction (model 2), ileum (model 1 and 2), and colon (model 1) by profiling nuclear and ribosome-bound RNAs with a modified RAISIN-seq (ribosomes and intact single-nucleus RNA-seq) protocol. Key transcription factors and differentially expressed genes (DEGs) driving cell state transitions and subtype proportion changes, as well as type 2–specific responses, were identified and functionally validated using in vivo Perturb-seq. (**B**) UMAP of 7640 enteric neurons colored by subcluster. Dashed circles indicate seven broad neuronal classes, which are annotated with subset-level identities. (**C**) Dot plots representing the relative mean expression of canonical enteric neuronal markers and newly identified cluster-specific marker genes ("Additional markers"). (**D**) Heatmap of the neuropeptide gene expression across each enteric neuronal subtype (left) and their cognate receptor genes in non-neuronal cell types in the intestine (right). Arrows indicate ligand-receptor pairs.

(i) *Chat*$^{+}$*Tac1*$^{+}$ putative excitatory motor neurons (PEMN), (ii) *Nos1*$^{+}$*Vip*$^{+}$ putative inhibitory motor neurons (PIMN), (iii) *Glp2r*$^{+}$ putative secretomotor/vasodilator neurons (PSVN), (iv) *Penk*$^{+}$ putative interneurons (PIN), and (v) *Chat*$^{+}$*Tac1*$^{-}$*Nos1*$^{-}$ putative sensory neurons (PSN), which could be further categorized in three distinct groups (Fig. 1C, left). Granular analysis revealed 19 distinct transcriptional clusters encompassing specific cellular states of broader cell types (Fig. 1, B and C). Clusters were further annotated according to cluster-specific genes (Fig. 1C, right; Materials and methods; and table S3), with benchmarking against published ENS atlases (*18*, *28*) (fig. S2, D and E). Motor neuron clusters were spread across related transcriptional states, whereas sensory neurons formed distinct subsets.

We next validated these cellular states by using marker genes for distinct neuronal subsets and assessing their expression with single-molecule

fluorescent in situ hybridization (smFISH) and immunofluorescent staining, colocalizing signals with class III β-tubulin (TUBB3) because some markers are also expressed by non-neuronal cells such as secretin receptor (SCTR) in intestinal crypts (fig. S3, A, B, D, and E). To investigate spatial organization, we quantified the localization of these markers along the mucosa-serosa axis and measured distances between neurons and the crypt base (fig. S3, B and C). Several neuronal subsets were categorized as myenteric plexus–enriched (PIMN *Aldh1a3*-hi, PIMN *Bglap*-hi, PSN *Grp*-hi, and PINs) or submucosal plexus-enriched (PSN *Piezo2*-hi and PSN *Vipr2*-hi), whereas some were found in both plexuses. PEMN *Fbxw15*-hi cells displayed a spread along the submucosa-serosa axis. Lastly, we validated the expression of newly identified marker genes—including EGFR (PIN), SCG2 (PEMN), MC4R (PSN), EPHA8 (PIMN), MOXD1 (PSVN), and BGLAP (PIMN)—through whole-mount staining of both submucosal and myenteric plexuses (fig. S3, D and E). To functionally characterize ENS clusters, we also analyzed the expression of genes from distinct gene families and biological processes, such as cytokine and chemokine signaling pathway genes; ion channels and transporters involved in mechanosensation, thermal sensitivity and pruriception; G protein–coupled family receptors (GPCRs); and transcription factors (TFs) (fig. S4). Our sequencing approach effectively captured lowly expressed genes and highlighted the selective expression of many GPCRs across distinct populations of sensory neurons.

Neuronal communication networks in the intestine

We reasoned that by mapping the expression profiles of neuroactive ligands and their corresponding receptors, we could construct a network of potential cell communication pathways between intestinal neurons and non-neuronal cells. To that end, we first created a pan-intestine non-neuronal atlas by integrating, clustering and annotating publicly available and newly generated single-cell transcriptomes of relevant mouse models (table S4, fig. S5, and Materials and methods).

Analyzing the expression of known neuropeptides and biosynthetic enzymes revealed putative circuits (Fig. 1D), for example, *Sst*-expressing PSN *Grp*/*Vipr2*-hi cells may communicate with *Sstr* family-expressing cells in the epithelial barrier, such as enteroendocrine cells (EECs) and goblet cells, and with vasculature, specifically pericytes and blood endothelial cells. Likewise, *Cck*-expressing PSN *Mc4r-hi* and *Piezo2*-hi neurons may cross-talk with *Muc6*$^{+}$ Brunner's gland cells localized in the duodenum, as well as *Nog*$^{+}$ interstitial cells of Cajal (ICC), through *Cckbr*. Whereas cells in the epithelial barrier expressed higher amounts of *Vipr1*, stromal cells and immune lymphocytic cells were enriched for *Vipr2*, suggesting compartment-specific encoding of the effects of vasoactive intestinal peptide (*Vip*) from PIMNs and PSVN *Npr1*-hi neurons (Fig. 1D). PIMNs are also canonically identified by high expression of *Nos1*, which encodes an enzyme involved in nitric oxide synthesis and is primarily detected by receptors enriched in stromal cells (fig. S6, A and B). PSN neurons were generally enriched for enzymes involved in dopamine synthesis: *Dbh*, *Ddc*, and *Th*. Dopamine receptors were expressed in goblet epithelial cells and were highly enriched in the non-enterochromaffin enteroendocrine cells (EEC 2), suggesting a dopamine-specific ENS-EEC cross-talk (fig. S6, C and D). Overall, we observed that the highest number of receptors for neuroactive molecules were expressed in the enterochromaffin cells (EEC 1) that synthesize serotonin (fig. S6E), suggesting extensive cross-talk between serotonin-producing cells and the ENS (*29*). Taken together, these data provide a comprehensive view of the cellular landscape of the ENS and elucidate potential cell circuits responsible for the cohesive functioning of the gastrointestinal system.

Regional imprinting revealed by shifts in cellular states of ENS

The segments of the intestine are distinguished by their distinctive anatomy, microbial contents, and physiological functions (*26*). Focusing first on SPF controls, we performed principal components analysis (PCA) of the pseudobulk transcriptome profiles (Materials and methods), revealing a distinct clustering of colonic relative to small intestine neurons (Fig. 2A).

We implemented two approaches to identify ENS cell types that are differentially abundant across regions. First, we used an embedding density-based approach to quantify the distribution of cell states in the uniform manifold approximation and projection (UMAP) embeddings across regions. Second, we used a *k*-nearest neighbor approach to create a graph of aggregated cellular neighborhoods, rather than using the predefined hard clusters, and assessed the differential abundance of the cellular neighborhoods (Materials and methods). We directly validated the observed changes (absolute cell counts versus percentage of neuronal subsets across intestinal regions, reflecting subset density and subset composition, respectively) by confocal imaging using immunostaining in independent animals. Within PEMNs, we observed that *Fbxw15*-hi and *Gda*-hi cells were enriched in the small intestine, whereas *Scg2*-hi cells were primarily found in the colon (Fig. 2, B and C, and fig. S7, A and B). Likewise, within PIMNs, we observed a shift from the *Aldha13*-hi state enriched in the small intestine to the *Bglap*-hi state in the colon (Fig. 2, B and C, and fig.S7, A and B). Specific neuronal subsets such as PSN *Grp*-hi and PIN *Penk*-hi cells were also enriched in the colon, whereas PSN *Mc4r*-hi, PSN *Nmu*-hi, and PSVN *Gal*-hi cells were enriched in the small intestine. Overall, these analyses revealed region-specific ENS–cell type enrichments and shifts in cellular states along the length of the intestine.

Next, we performed differential expression (DE) analysis to identify regionally imprinted genes after regressing the effect of cell type, sex, and technical covariates (Materials and methods). As a confirmation, we observed a colonic enrichment of *Ahr*, *Ano5*, *Pou3f3*, *Col25a1*, and *Pde1c*, which has been reported in previous studies (*8*, *9*) (Fig. 2D and table S5). Multiple immune-related receptors were regionally enriched, suggesting potential region-specific neuroimmune interactions. For example, colonic neurons were enriched for receptors for type 2 cytokines (*Il4ra* and *Il13ra1*) and type I interferon (*Ifnar1*). By contrast, the duodenum-jejunum (DJ) region was enriched for receptors for complement binding (*Cd46*) and interferon-gamma (*Ifngr2*) (Fig. 2D). Gene ontology (GO) analysis revealed distinct pathway enrichments across regions (Fig. 2E): DJ exhibited enrichment in RNA splicing and chromatin regulation; ileum was associated with neuronal development, Wnt signaling, and circadian rhythm (such as *Per1* in Fig. 2D); and colon genes were primarily linked to metabolic pathways (including sterol biosynthetic and glycolytic processes). Evaluating region-specific TF regulon activity (Materials and methods) unveiled higher activity of *Fos* and *Jun* signaling in the colon. Conversely, *Gata3* activity was enriched in the DJ and ileum (fig. S7C).

Microbiome signals distinctly regulate ENS composition in the ileum and colon

We next explored how the ENS responds to changing microbial colonization. First, we investigated how enteric neurons in conventionally raised mice (SPF) differ from those in GF and recolonized mice (Materials and methods). PCA analysis and hierarchical clustering of pseudobulked transcriptome profiles revealed that neurons in SPF-born mice separated from all GF-born mice (GF, BT, and ASF) regardless of intestinal segment (Fig. 3A and fig. S8A). Given these observations suggesting a dominant impact of colonization with a complex microbiome from birth, we first grouped the mice into two categories: animals born GF and born SPF.

Composition analysis of born GF and born SPF mice uncovered region-specific shifts. For instance, within PIMNs in the small intestine, *Bglap*-hi cells were enriched in born SPF mice, whereas *Aldh1a3*-hi cells were enriched in the born GF mice; this finding was also validated using confocal imaging (Fig. 3, B and C, and fig. S8B). Notably, we observed a reversal of that trend in the colon: *Bglap*-hi cells were enriched in the born GF condition and *Aldh1a3*-hi cells in born SPF.

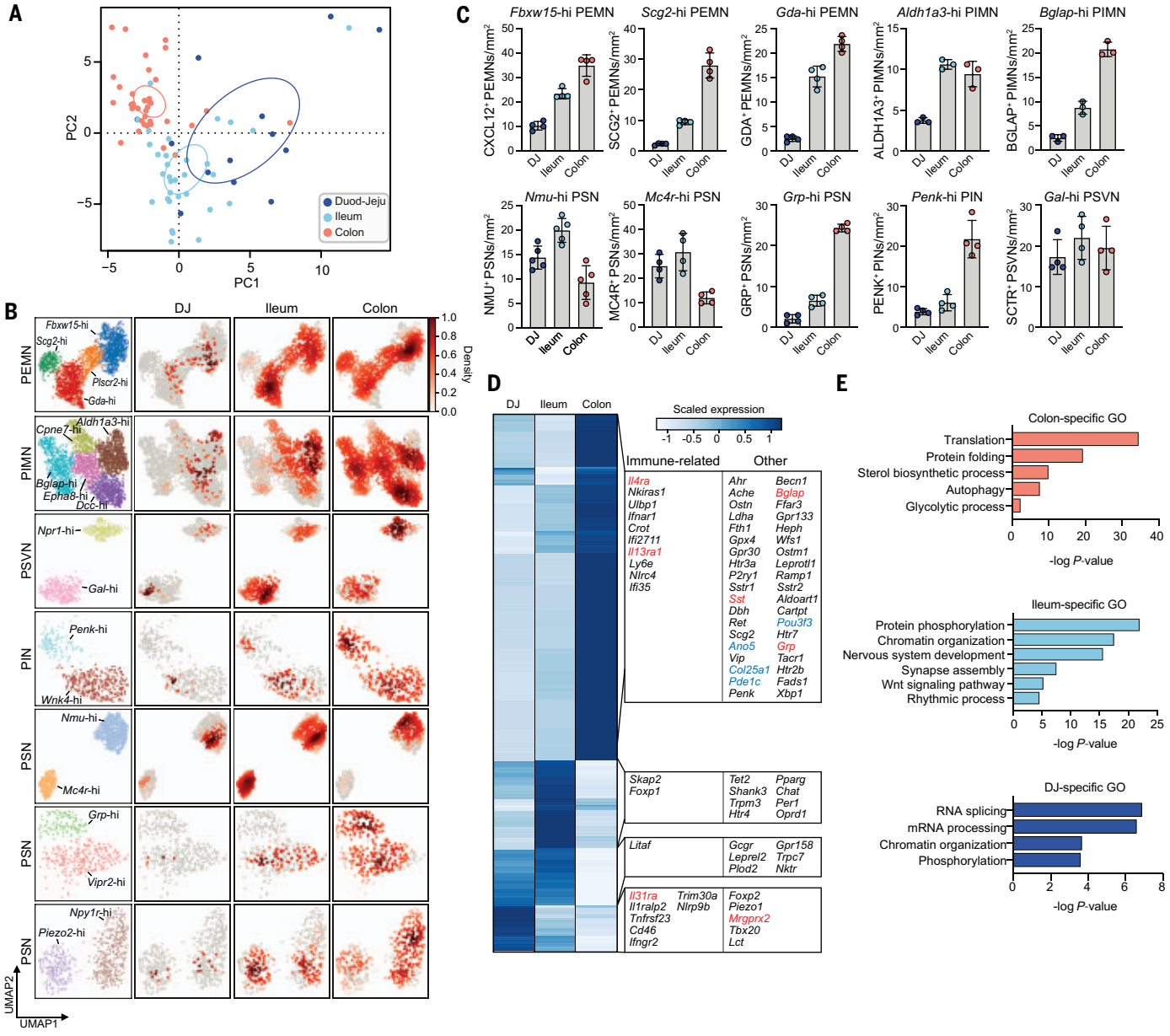


Fig. 2. Regional imprinting revealed by shifts in cellular states of ENS. (**A**) PCA plot showing the similarity of enteric neurons across duodenum-jejunum (DJ), ileal, and colonic regions. Each dot represents the pseudobulk transcriptome profile of all cells from an individual mouse and their respective tissue site. (**B**) Embedding densities of cells in each broad neuronal cell type across the DJ, ileal, and colonic regions of the intestine. Each broad neuronal cell type was subset from the complete atlas, and the original UMAP embeddings (shown with subset labels in the leftmost column) were used to compute the density. (**C**) Quantification of ENS subsets and their layer-specific distribution (number/mm^2) using cross-sectional immunofluorescence (IF) confocal images of the DJ, ileal, and colonic regions in SPF wild-type mice. (**D**) Heatmap showing region-specific ENS genes across the DJ, ileal, and colonic regions. Only genes with adjusted $P < 0.05$ and average $\log_2$ fold change > 0.25 were included. Selected genes related to immune functions and other genes associated with neuronal functions and enriched biological processes are displayed. Previously reported region-specific genes are highlighted in blue; key genes relevant to this study are highlighted in red. (**E**) Gene ontology (GO) biological process analyses of region-specific ENS genes across the DJ, ileal, and colonic regions, performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID).

Such a reversal in region-specific shifts was also observed in other neuronal subtypes: *Scg2*-hi PEMNs were expanded in born SPF ileum and born GF colon, and, conversely, *Fbxw15*-hi PEMNs were increased in born GF ileum and born SPF colon (Fig. 3, B and C, and fig. S8B). We also observed an increase in PSN *Grp*-hi cells specifically in born GF colon and additional region-dependent but not colonization-dependent changes, which we validated with immunofluorescence (Fig. 3B and fig. S8C).

Next, we performed DE analysis to identify gene programs induced or repressed by microbial colonization, again regressing out the effects of cell type, sex, and technical covariates (Materials and methods). We observed a greater overlap ($n = 215$) of the differentially expressed genes (DEGs) induced in born GF colon with the DEGs induced in born SPF ileum (Fig. 3D and table S6). Moreover, genes induced in the born GF colon were conspicuously suppressed in the born GF ileum (Fig. 3, E and F, and table S6). Likewise, the share of DEGs induced in

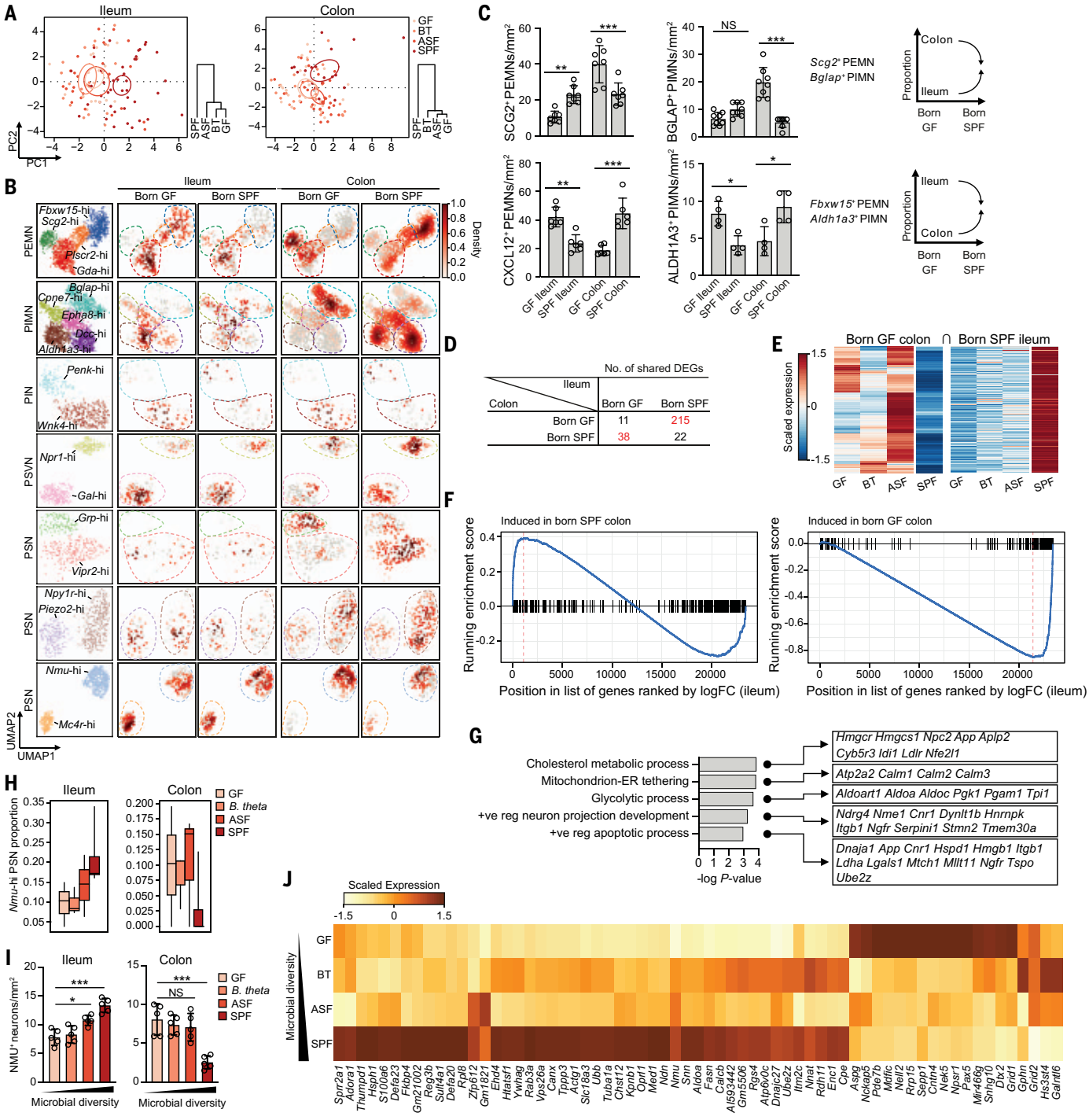


Fig. 3. Microbiome signals distinctly regulate ENS composition in the ileum and colon. (**A**) PCA plot showing the similarity of ENS cells in GF mice with or without colonization with *B. theta* (BT) or ASF, compared with SPF mice, in ileum and colon. Each dot represents the pseudobulk transcriptome profile of all cells of a given cell type from an individual mouse and the respective condition. Dendrograms on the right of the panels show hierarchical clustering of each condition to assess the effect of microbial colonization. Dendrogram analysis was performed on the top 100 PCs. (**B**) Embedding densities of cells within each broad neuronal cell type from mice born GF (combined GF, BT, and ASF) or born SPF in ileum and colon. Each broad neuronal cell type was subset from the complete atlas, and the original UMAP embeddings (shown with subset labels in the leftmost column) were used to compute the density. (**C**) Quantification of PEMN and PIMN subset surface densities (neurons/mm^2) across GF, BT, ASF, and SPF conditions using whole-mount muscularis propria confocal IF images (*P* values: SCG2$^+$ PEMN = 0.0097, 0.0006; BGLAP$^+$ PIMN = 0.1919, <0.0001; CXCL12$^+$ PEMN = 0.0014, <0.0001; ALDH1A3$^+$ PIMN = 0.0245, 0.0146). Schematic diagrams on the right illustrate modes of proportional changes (indicated by arrows) from born GF to born SPF in ileum or colon (see also fig. S8C). NS, not significant. (**D**) Table showing the number of shared DEGs enriched across each indicated comparison of regions and conditions. Only genes with adjusted $P < 0.05$ and $\log_2$ fold change (FC) > 0 were considered for each class. (**E**) Heatmaps showing pan-neuronal scaled mean expression of developmentally imprinted genes. (Left) Expression in the colon for genes induced in both the born GF colon and the born SPF ileum. (Right) Expression of the same genes across conditions in ileum. (**F**) Gene set enrichment analysis plot showing opposite regional encoding of microbial colonization. (Left) Enrichment plot of genes significantly induced in born SPF colon (adjusted $P < 0.05$) ranked by the logFC in ileum. (Right) Enrichment plot of genes significantly induced in the born GF colon (adjusted $P < 0.05$), ranked by the logFC in ileum. A negative FC on the *y* axis indicates enrichment in born SPF

ileum, whereas a positive log FC means enrichment in born GF ileum. (**G**) Gene ontology (GO) analysis of 215 DEGs shared between born GF colon and born SPF ileum. ER, endoplasmic reticulum; +ve reg, positive regulation. (**H** and **I**) Proportion plots and quantification of neurons showing changes in the proportion of PSN *Nmu*-hi cells in ileum or colon with increasing microbial load and diversity, from GF to BT to ASF to SPF. Proportions were estimated using the scCODA package (H). Quantified number of NMU^+TUBB3^+ neurons in ileum or colon across GF and bacterial spectrum (BT, ASF, SPF, $n = 5$ mice per group) is shown (I). (**J**) Heatmap showing the mean scaled expression of genes associated with microbial load (density and diversity) in ileal PSN *Nmu*-hi cells.

born GF ileum was higher in born SPF colon ($n = 38$) than in born GF colon ($n = 11$) (Fig. 3D and table S6). Functional enrichment analysis of the 215 DEGs shared between born GF colon and born SPF ileum (Fig. 3, D and E) highlighted biological processes enriched for cholesterol metabolism and glycolysis, as well as positive regulation of neuron projection and apoptosis (Fig. 3G). These results suggest that microbial signals present during development or early life induce broadly opposite shifts in cellular states and gene programs observed in the small intestine as compared with the large intestine.

We further explored our data to identify microbial diversity–specific changes from GF to BT to ASF to SPF. We observed that the proportion of ileal PSN *Nmu*-hi cells increased alongside microbial colonization, whereas colonic *Nmu*-hi PSN were not altered outside of SPF animals (Fig. 3, H and I, and fig. S8D). Within ileal PSN *Nmu*-hi cells, we further used linear regression modeling to identify genes responsive to increasing microbial load and diversity (Materials and methods and Fig. 3J). We observed that genes involved in neuronal development (*Ywhag*, *Sult4a1*, *Nnat*, *Rgs4*, *Ndn*, *Oprl1*, and *Cpe*) and the neuropeptides Nmu and Calcb were induced with increasing microbial load.

To further investigate the influence of microbial complexity and colonization timing on neurons, we conducted additional experiments comparing GF mice colonized as adults with ASF or SPF microbiomes to mice born and raised under ASF or SPF conditions. As noted, short-term ASF colonization had a minimal impact on ENS subsets (fig. S9), except for ileum *Nmu*-hi PSNs. By contrast, long-term colonization or more complex microbiomes led to broader shifts in ENS composition: Short-term SPF colonization exerted a strong influence on multiple motor neuron subsets and largely recapitulated the composition of born SPF animals for PEMN and PIMN states, whereas colonization with ASF microbes from birth showed only a partial rescue (fig. S9, A to D). These results suggest differential sensitivities of the sensory and motor compartments to microbial signals and an additive role of microbiome complexity and time of colonization on the ENS, with ileal *Nmu*-hi PSNs being distinctly responsive to even short-term colonization with a simple microbiome.

Role of Grp in modulating peristalsis

Gastrin-releasing peptide (Grp) is a neuropeptide expressed in the gastrointestinal tract whose primary role is stimulating gastrin secretion by G cells in the stomach to regulate food intake and digestion (*30*). We found that *Grp* was enriched in the colon, particularly in subsets of sensory neurons coexpressing *Sst* (PSN *Grp*-hi) or *Nmu* (PSN *Nmu*-hi) (Fig. 2C and Fig. 4, A to C). As in other cell types (Fig. 3), we observed an opposite regional effect of microbiota colonization on the number of Grp^+ neurons (Fig. 4D). Furthermore, we observed that Grp's receptor, *Grpr*, was enriched in Nog^+ ICC and goblet cells in our non-neuronal dataset (Fig. 4E). As Grp^+ neurons were modulated by microbiota, we investigated whether this modulation extended to ICC cell number and the expression of *Grpr*. The total number of ICCs, defined as $ANO1^+$, increased in GF versus SPF mice in ileum (Fig. 4, F and G). However, the $GRPR^+$ ICC population was elevated specifically in the colon of GF mice as assessed both by immunofluorescence (Fig. 4H) and single-cell genomics (Fig. 4, I and J). Other cell types in the colon remained unchanged between GF and SPF conditions (fig. S8, E and F), except for $Kcnn3^+$ fibroblasts (Fig. 4I and fig. S5C)—which likely represent the fibroblast subset that is associated with ICCs and smooth muscle cells (*31*)—and plasma cells, which are down-regulated in GF animals. Given that the GF mice are known to exhibit dysregulated peristalsis (*32*), our finding that PSN *Grp*-hi cells covary with GRPR-expressing ICCs suggests a potential function of Grp in regulating intestinal transit.

Pan-neuronal responses to type 2 inflammation

Type 2 immunity is associated with both protection from helminth infections and detrimental allergic responses. Neurons also play a major role in mediating defensive responses, such as diarrhea and vomiting to expel noxious stimuli. Pan-neuronal pseudobulk transcriptome profiles of type 2 responses to OVA sensitization (OVAs), OVA challenge (OVAc), and *H. polygyrus* infection (HP^+) revealed both shared genes (*Il31ra*, *Zfp599*, and *Vsig4*) (Fig. 5A and table S7) and pathways between OVAc and HP infection (fig. S10, A to D). A subset of genes was only induced in OVAs samples, including the chemokine receptor (*Darc*) (Fig. 5A), suggesting that enteric neurons may also be primed during sensitization. Consistent with previous studies of type 2 immunity, the antimicrobial peptide *Sprr2a2* was up-regulated in the OVAs, OVAc, and HP^+ conditions (*33*).

We next sought to identify mediators of ENS-immune interactions. Using our non-neuronal atlas, we identified multiple type 2 inflammation–induced ligands whose receptors were also expressed in ENS subsets. Specifically, in ILC2 cells we observed up-regulation in amphiregulin (*Areg*); CGRP (*Calca*); and cytokines *Il13*, *Il4*, and *Lif*, whose receptors were expressed in ENS subsets including PSN *Nmu*-hi cells—the only subset to express both *Il13ra1* and *Il4ra*. *Nmu*-expressing neurons were previously reported to interplay extensively with type 2 responses in the gut (*15*), and a recent study demonstrated that sensing of type 2 cytokines is essential for proper responses to parasites (*34*) (Fig. 5B). Mast cells up-regulated oncostatin M (*Osm*) and histidine decarboxylase (*Hdc*), with PSN *Mc4r*-hi cells the most likely responders to OSM and histamine as they expressed the highest amounts of *Osmr* and *Hrh1* (Fig. 5B), suggesting another potential neuroimmune circuit.

Adaptive cellular states of PSN *Nmu*-hi cells to Th2 inflammation

Compositional analyses of the single-nucleus data from OVAs, OVAc, and HP models did not show significant subset-level changes, but neighborhood analyses highlighted loss of specific populations of PSN *Nmu*-hi cells (fig. S10, E and G). We therefore performed confocal imaging and validated this reduction in PSN *Nmu*-hi cells in both OVAc and HP^+ samples (Fig. 5, C and D). This reduction could be due to overactivation-induced cell death—as evidenced by the down-regulation of antiapoptotic genes (*Bcl2l1* and *Mcl1*, as well as an enrichment in the negative regulation of the neuronal apoptotic process GO pathway) in OVAc (fig. S10H and table S8)—and up-regulation of pathways involved in positive regulation of the apoptotic process in HP^+ (fig. S10I and table S8). We also observed an induction of *Gpr158* in *Nmu*-hi PSNs in the OVAc model (Fig. 5E). To validate changes in *Gpr158* expression, we performed whole-mount staining of the muscularis propria from OVAc and HP^+ mice. *Gpr158* expression was found to be neuron-specific and up-regulated in response to type 2 inflammation (Fig. 5, F and G). *Gpr158* is highly expressed in enteric neurons (fig. S10, J and K) and was recently described as a metabotropic glycine receptor, but its biology is also influenced by osteocalcin (encoded by *Bglap*), with known roles in stress responses (*35*–*37*).

To further assess whether type 2 cytokines are key drivers of the observed transcriptomic (Fig. 5A) and PSN *Nmu*-hi–specific (Fig. 5, C and E) changes, we performed additional experiments involving

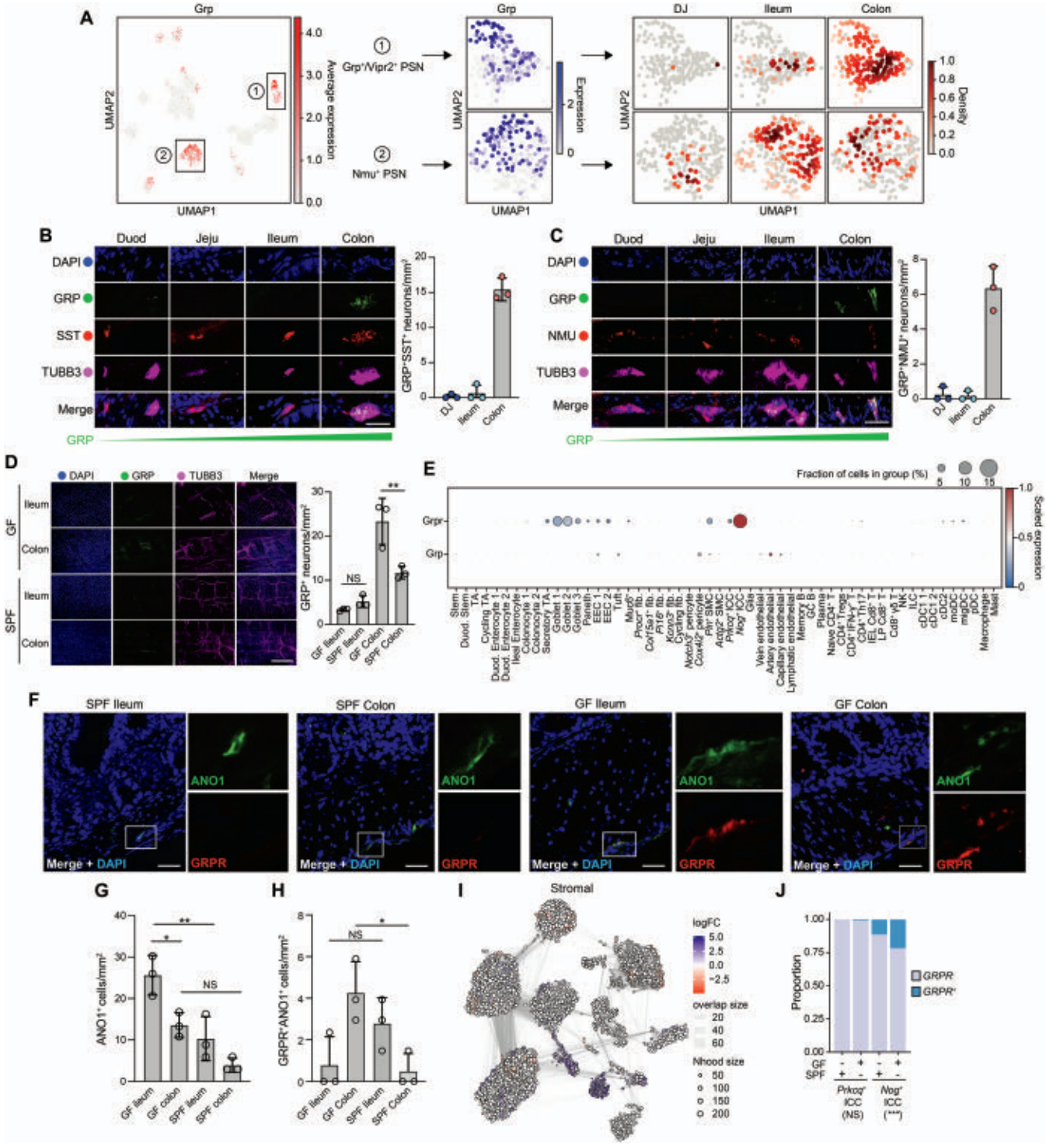


Fig. 4. Grp-hi sensory neurons and Grpr-expressing ICCs are expanded in germ-free colons. (**A**) (Left) UMAP of *Grp* expression in the overall atlas, showing enrichment in PSN *Grp*-hi and PSN *Vipr2*-hi cells (Grp-Vipr2 PSN) (1) and PSN *Nmu*-hi cells (2). (Middle) *Grp* expression in the enriched cell types. (Right) Embedding densities of the enriched cell types across DJ, ileal, and colonic regions. Each broad neuronal cell type was subset from the complete atlas, and the original UMAP embeddings were visualized and used to compute the density. (**B** and **C**) Representative confocal IF images showing increased GRP expression (green) in PSN *Nmu*-hi (NMU, red; TUBB3, magenta) and PSN *Vipr2*-hi cells (SST, red; TUBB3, magenta) across DJ, ileal, and colonic regions. Quantification of GRP$^+$ neurons (number/mm^2) from cross-sectional confocal IF images is shown on the right ($n = 3$ per group). Scale bars, 25 μm. (**D**) Representative confocal IF images showing GRP$^+$ neurons (GRP, green; TUBB3, magenta) in muscularis propria across indicated regions and conditions. Quantification of GRP$^+$ neurons (neurons/mm^2) from whole-mount confocal IF images is shown on the right ($n = 3$ per group; *P* values: GRP$^+$ neurons = 0.0043). Scale bars, 100 μm. (**E**) Dot plots showing relative expression of *Grpr* and *Grp* in all non-neuronal cell types in mouse colon. (**F**) Representative confocal IF images showing ANO1$^+$ and GRPR$^+$ cells across the indicated regions and conditions. DAPI was used as a nuclear counterstain. Scale bars, 25 μm. (**G**) Quantification of ANO1$^+$ICC cells (cells/mm^2) from cross-sectional confocal IF images across indicated regions and conditions ($n = 3$ per group; *P* values for ANO1$^+$ cells: GF ileum versus colon = 0.0248; GF versus SPF ileum = 0.0066; GF versus SPF colon = 0.0681). (**H**) Quantification of GRPR$^+$ANO1$^+$ ICC cells (cells/mm^2) from cross-sectional confocal IF images across indicated

regions and conditions ($n = 3$ per group; P values for GRPR$^+$ANO1$^+$ cells: GF versus SPF ileum = 0.2849; GF versus SPF colon = 0.0265). (**I**) Neighborhood plots obtained from the miloR package showing differential abundance of non-neuronal stromal cell types between born GF and born SPF mice. Nodes represent neighborhoods, colored by their log FC. Only neighborhoods with significant differences [false discovery rate (FDR) < 10%] are colored (GF enriched neighborhoods are purple, and GF depleted neighborhoods are red); not-significant neighborhoods are colored white. Node sizes correspond to the number of cells in each neighborhood (Nhood). Graph edges indicate shared cell numbers between neighborhoods. The node layout is determined by the position of the neighborhood index cell in the UMAP in fig. S5C. (**J**) Proportion of *Grpr*+ cells within the *Prkcq*$^+$ ICC and *Nog*$^+$ ICC subsets in GF and SPF colon. P values from χ^2 test are reported.

IL-4Rα (shared subunit used by both IL-4 and IL-13 receptors) blockade and a systemic anaphylaxis model, characterized by elevated type 2 cytokines and OVA-specific IgE (fig. S11, A to G). We identified a set of type 2 inflammation–specific genes from previously characterized Th2-associated transcripts in sorted Uchl1$^+$ neurons (Fig. 5A), which were reversed upon IL-4Rα blockade, as exemplified by the down-regulation of *Gpr158* (fig. S11H and Fig. 5F). These same genes were up-regulated during anaphylactic shock (fig. S11I). In addition, blocking IL-4Rα signaling restored the proportion of PSN *Nmu*-hi cells reduced by OVAc (Fig. 5C). These findings demonstrate that type 2 cytokine signaling is a key regulator of both transcriptional programs and the abundance of Nmu-hi PSNs during allergic inflammation.

We next leveraged our atlas to investigate other potential signals involved in Nmu-hi PSN plasticity (*38*–*49*). We noted an up-regulation of key leukotriene synthesis enzymes in mast and tuft cells (Fig. 5H). The leukotriene receptor CysLTR2 is specifically expressed in PSN *Nmu*-hi cells and induced during Th2 inflammation (Fig. 5, H and I, and fig. S4D). To functionally validate the role of receptors for IL-13 and leukotrienes in PSN *Nmu*-hi cells, we performed Ca^{2+} imaging using the calcium indicator Fluo-4 on in vitro cultured enteric neurons from adult mouse intestine stimulated with IL-13 and LTC_4. Our results confirmed that NMU$^+$CHAT$^+$NOS1$^-$ neurons constituted the majority of neurons that functionally respond to IL-13 and LTC_4 (Fig. 5J) and vice versa.

Taken together, our findings suggest intrinsic gene programs and extrinsic cell circuits that are either specific to or concurrently engaged by OVA-induced allergy and HP infection. Specifically, we highlight that PSN *Nmu*-hi cells exhibit high plasticity driven by signaling from type 2 cytokines and potentially leukotrienes.

Neuronal differentiation potentially explains shifts in PIMN cellular states

We also observed subtype-specific changes within PIMNs in response to OVAc and HP infection, specifically the expansion of PIMN *Dcc*-hi and *Aldh1a3*-hi cells alongside a reduction of PIMN *Bglap*-hi cells (Fig. 5, K to M, and fig. S10, F and G). These alterations were reversed by IL-4Rα blockade, indicating their dependence on type 2 cytokine signaling (Fig. 5, L and M). These modulations aligned with the shifts induced by microbial colonization, including, for example, from PIMN *Bglap*-hi to *Aldh1a3*-hi/*Dcc*-hi cells (Fig. 3). RNA velocity analysis of PIMN cells (Materials and methods) showed a much higher ratio of aggregate spliced and unspliced reads in *Bglap*-hi cells compared with other PIMNs, suggesting the existence of a differentiation trajectory (Fig. 5N). Velocity-inferred latent pseudotime analysis suggested that PIMN *Bglap*-hi cells were at a relatively "immature" state along the latent time, whereas PIMN *Cpne7*-hi and *Epha*-hi cells coincided with an intermediate stage, and lastly that PIMN *Dcc*-hi and *Aldh1a3*-hi cells coincided with a "mature" stage (Fig. 5O). *Nos1* expression was lowest in *Bglap*$^+$ and gradually increased along this trajectory, whereas *Tubb3* expression was highest in *Bglap*$^+$ PIMNs (Fig. 5P). Altogether, these analyses support a differentiation trajectory within PIMNs that can be modulated by environmental signals.

In vivo Perturb-seq reveals key regulators of PIMN differentiation trajectory and neuronal programs

We next set out to identify regulators of neuronal programs in vivo, considering three key families of genes as potential targets (table S9): first, genes that are induced by microbiota signals and type 2 signals in PSN *Nmu*-hi cells, which closely interact with immune cells (Fig. 3J, Fig. 5E, and fig. S12B); second, genes that are induced in type 2 inflammation across all neurons (Fig. 5A and fig. S12C); and third, genes whose expression changed across the PIMN differentiation trajectory (Fig. 5, N and O, and fig. S12D).

AAV-based gene delivery and CRISPR screening have proven to be efficient methods for studying neuronal function (*50*, *51*). We generated separate pools of guide RNAs (gRNAs) for the different classes of genes discussed above and introduced them in Cas9-expressing animals by injecting AAV-PHP.S—a capsid that efficiently infects peripheral neurons (*52*)—vectors encoding these gRNAs upstream of a tdTomato-KASH reporter, flanked by piggyBac transposon elements (*53*) (Fig. 6A). Four weeks after injection, we isolated tdTomato$^+$ nuclei from submucosal and myenteric plexuses (fig. S12A) and subjected them to snRNA-seq.

Dimensionality reduction and clustering of the in vivo Perturb-seq data after quality control yielded 30,941 cells (Materials and methods) separated across clusters in general agreement with our SMART-seq2–based atlas (Fig. 6B and fig. S12, E and F). We did observe some differences in the clustering, likely due to the reduced depth of snRNA-seq compared with SMART-seq2 or differences in nuclei isolation protocols; notably, PIMN *Cpne7*-hi, PEMN *Plscr2*-hi, and PSN *Grp*-hi cells were not identified as separate subsets, and PSN *Nmu*-hi cells were recovered in lower numbers (Fig. 6B and fig. S12F). Yet the similarity of the two atlases confirms the overall agreement between *Uchl1* reporter– and AAV-based approaches (fig. S12F, middle and right). Around 49% of cells had one or multiple gRNAs, with 9298 cells assigned a single gRNA perturbation, with guide recovery generally uniform across clusters (fig. S12F, left). We also included a further 1178 cells with multiple gRNAs but a single dominant gRNA (Materials and methods) for a total of 10,476 cells and a median of 44 cells per gRNA.

Next, we sought to identify perturbations that could induce a change in ENS composition (fig. S12G). To identify perturbations that induce a significant change while correcting for differences in cell and guide, we used a sample size–dependent permutation approach to generate a background distribution of compositions for a given number of cells and assessed the significance of the distance of the guide-associated cells in reference to this background in principal component (PC) space (Materials and methods). Using this approach, we identified that gRNAs targeting *Asic4*, *Mafg*, *Ncoa1*, *Npas2*, and *Zfp804b* induced a significant change ($P < 0.05$) in the overall neuronal cell type composition (Fig. 6, C and D). Inspection of cell type composition for the target gene perturbations suggest that deletion of *Asic4*, *Mafg*, and *Ncoa1* leads to an expansion in PIMN *Epha8*-hi cells relative to the nontargeting controls (Fig. 6D).

To identify potential regulators of neuronal type 2 responses, we first derived a conserved type 2 response–associated gene program signature at the pan-neuronal level (Materials and methods) and used up- and down-regulated genes to compute two distinct neuronal type 2 response scores (neuronal type 2–induced and –repressed). We next modeled the effect of genetic perturbations on these programs by fitting a multivariate linear regression adjusting for technical covariates (Materials and methods). Among the set of perturbations impacting pan-neuronal type 2–related genes, the deletion of *Zfp804a*, *Gpr98*, or *Pgk1* decreased the type 2–induced program, whereas deletion of *Sprr2a2* increased the type 2–induced program (Fig. 6E). Our perturbation results suggest that deletion of these genes restrains the induction of the type 2 immunity–induced response in enteric neurons.

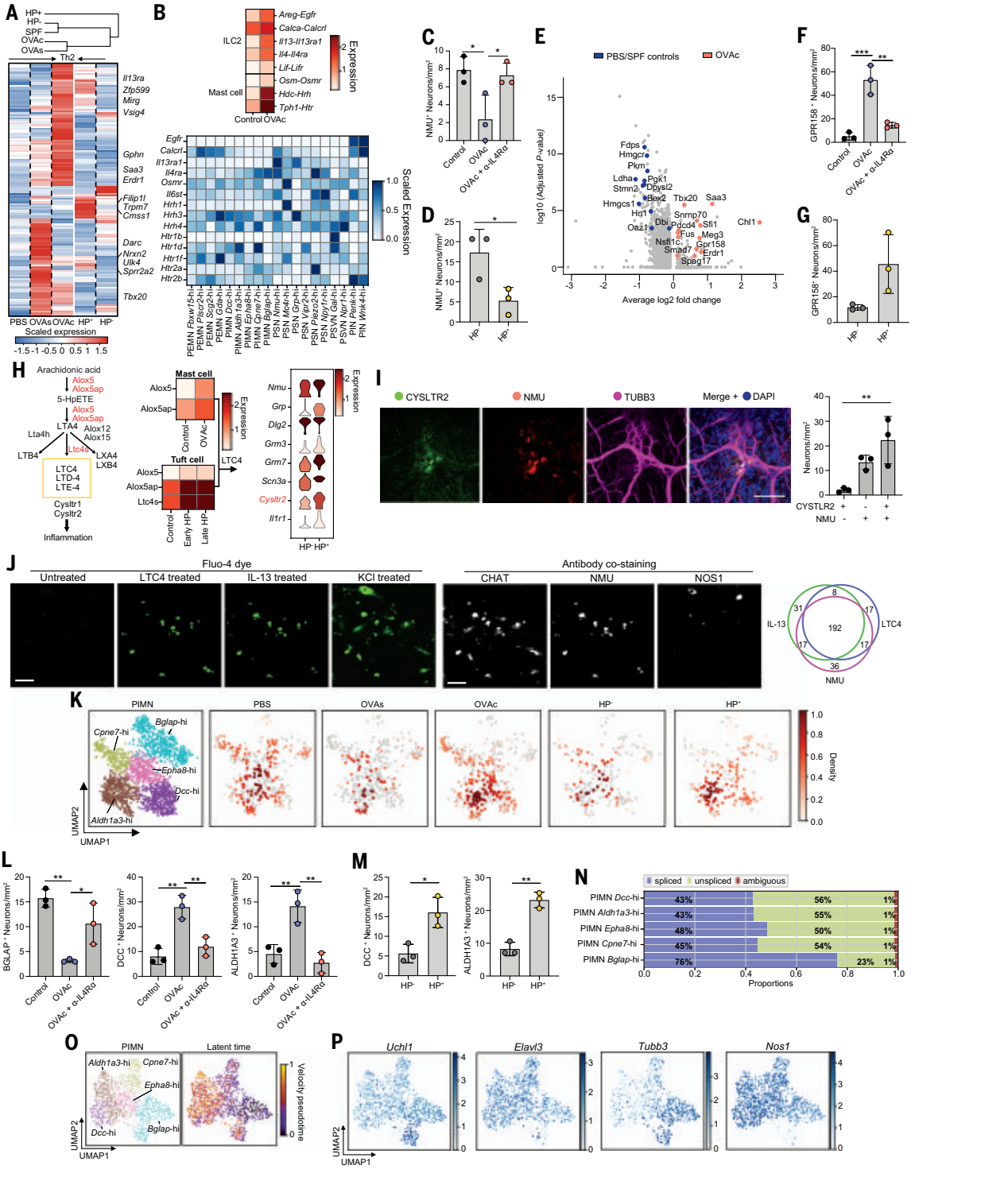


Fig. 5. Responses of *Nmu*-hi PSN and PIMNs to type 2 inflammation. (**A**) Heatmap showing the relative mean expression of pan-neuron shared and distinct DEGs in ileum across type 2 inflammation mouse models: OVA sensitization (OVAs) or OVA challenge after sensitization (OVAc) [versus control, including phosphate-buffered saline (PBS)-treated and SPF mice] and helminth infection (HP^+ versus HP^-). Dendrograms above the heatmap show hierarchical clustering of each condition to assess study type 2 inflammation effects. Dendrogram analysis was performed with the top 100 PCs. (**B**) Mean expression of selected food allergy–induced ligand genes in ILC2 and mast cells (left panels) and their cognate receptor genes expressed in ENS subsets (right panels). Differential expression analysis was performed using the Wilcoxon test in the non-neuronal atlas between OVA-treated and control cells. Only ligand genes with adjusted $P < 0.05$ are shown. (**C** and **D**) Quantification of NMU^+TUBB3^+ neurons/mm^2 in ileum from whole-mount muscularis propria confocal IF images under the following conditions: (C) control, OVAc, and OVAc + α-IL4Rα (type 2 cytokine blockade) (*P* values: Control versus OVAc = 0.0310; OVAc versus OVAc + α-IL4Rα = 0.0485) or (D) HP^+ versus HP^- ($n = 3$ per group; $P = 0.0242$). (**E**) Volcano plots showing DEGs in PSN *Nmu*-hi cells between OVAc and control mice (PBS-treated + SPF mice). Genes with an adjusted $P < 0.2$ are highlighted. (**F** and **G**) Quantification of $GPR158^+$ $TUBB3^+$ neurons/mm^2 in ileum from whole-mount muscularis propria confocal IF images under the following conditions: (F) control, OVAc, and OVAc + α-IL4Rα (*P* values: Control versus OVAc = 0.0006; OVAc versus OVAc + α-IL4Rα = 0.0018) or (G) HP^+ versus HP^- ($n = 3$ per group). (**H**) (Left) Biosynthesis pathway of leukotrienes with key enzymes highlighted in red. (Center) Heatmaps

showing induction of key leukotriene biosynthesis enzymes in mast cells and tuft cells during type 2 inflammation. "Early HP" and "Late HP" represent combined data from days 3 and 4 or days 10 and 14 of helminth infection, respectively. (Right) Violin plots showing the expression of Th2-associated genes (*Nmu* and *Grp*) and genes involved in pain and itch responses (*Dig2*, *Grm3*, *Grm7*, *Scn3a*, *Cysltr2*, and *Il1r1*) in ileal *Nmu*-hi PSN, in HP^+ versus HP^- mice. (**I**) Representative confocal IF images of $CysLTR2^+NMU^+$ neurons in ileum (CysLTR2, green; NMU, red; TUBB3, magenta). DAPI was used as a counterstain. Scale bars, 25 μm. Quantification of $CysLTR2^+NMU^+$ neurons/mm^2 from whole-mount muscularis propria confocal IF images is shown on the right ($n = 3$; $P = 0.0124$). (**J**) Representative Ca^{2+} imaging of NMU^+ neurons, untreated or sequentially treated with IL-13 (ligand for IL13RA1), leukotriene LTC_4 (ligand for CysLTR2), and KCl (a pan-neuronal calcium flux inducer). Scale bars, 25 μm. (Rightmost panel) Venn diagram showing the number of overlapping IL-13–responsive cells, LTC4-responsive cells, and NMU-expressing cells in 12 quantified Ca^{2+} imaging areas ($n = 3$ mice, four quantified areas each repeat) covering a total of 318 neurons; 192 (60%) are positive for all three and 234 (73.6%) for at least two out of three readouts. (**K**) Embedding densities of small intestinal PIMN cells (ileum and DJ) across the spectrum of type 2 conditions. The original atlas was subset to PIMNs, and the UMAP embeddings (leftmost panel, labeled by subset) were used to compute and visualize density across experimental conditions. (**L** and **M**) Quantification of $BGLAP^+$, DCC^+, or $ALDH1A3^+$ neurons from whole-mount muscularis propria under the following conditions is shown: (L) control, OVAc, and OVAc + α-IL4Rα (*P* values: $BGLAP^+ = 0.0025$, 0.0294; $DCC^+ = 0.0021$, 0.0065; $ALDH1A3^+ = 0.0077$, 0.0033) or (M) HP^+ versus HP^- (*P* values: $DCC^+ = 0.0159$; $ALDH1A3^+ = 0.0012$) ($n = 3$ per group). (**N**) Proportion of aggregate splicing reads across PIMN subsets, showing that PIMN *Bglap*-hi cells have a higher proportion of spliced reads, suggesting a progenitor-like cellular state. (**O**) Estimated latent velocity pseudotime of PIMNs in the UMAP embedding, plotted alongside the same embedding with subsets labels to show the differentiation trajectory of PIMNs. UMAP embeddings were recomputed on the subset of 1298 PIMNs that were used for velocity estimation. (**P**) Expression plots of markers of neuronal identity (*Uchl1*, *Elavl3*, and *Tubb3*) and PIMN (*Nos1*) onto the UMAP projection.

We utilized a similar method to identify putative regulators of PIMN differentiation by first using a consensus non-negative matrix factorization (cNMF)–based approach (*54*) to derive gene modules within PIMNs, recovering six distinct gene programs (Fig. 6, F and G). We then modeled the effects of genetic perturbations using the nontargeting controls as the reference (Fig. 6H and fig. S12H). This analysis recovered *Ncoa1*, *Mafg*, and *Zfp804b* deletions as having an effect on PIMN cell states, lending support to the conclusions of the composition analysis (Fig. 6, C and D). Specifically, we observed an upregulation of the PIMN *Epha8*-hi cell–specific module 1 and an inverse effect on the PIMN *Aldh1a3*-hi module 4, after *Ncoa1* deletion (Fig. 6, D, G, and H). Similarly, *Edf1* deletion had a positive effect on the PIMN *Bglap*-hi cell–specific module 6 (Fig. 6, G and H), which is consistent with the observed expansion of PIMN *Bglap*-hi cells (Fig. 6I). Given that Edf1 is known to be a regulator of endothelial differentiation (*55*) and that its highest expression is in PIMN *Bglap*-hi (fig. S12D), it is plausible that it could also play a role in PIMN neuronal differentiation. Conversely, deletion of *Mitf* (expressed in PIMN *Cpne7*/*Aldh1a3*/*Dcc*-hi), *Setdb1* (expressed in PIMN *Cpne7*-hi), or *Zfp804b* (expressed in PIMN *Bglap*-hi) had a positive effect on the PIMN *Dcc*-hi cell–specific gene module 3 along with PIMN *Dcc*-hi expansion (Fig. 6, G and H). Taken together, this work suggests multiple modulators of PIMN states and demonstrates the utility of in vivo Perturb-seq screens to study genetic regulation in the ENS at medium throughput.

Next, we aimed to validate the effects and phenotypic impact of individual genes by leveraging the relatively high transduction efficiency of AAV-PHP.S (78.35% ± 1.51%, fig. S13) (*56*, *57*). By deleting a single gene in >75% of neurons, we expected that infection might lead to measurable phenotypes, such as modulation of gut transit. We infected animals with AAVs expressing control guides or guides for six genes with potential relevance to PIMN biology (*Mafg*, *Edf1*, *Mitf*, *Zfp804b*, *Tbx20*, and *Asic4*). We found that deletion of *Edf1* (predicted as a regulator of $Bglap^+$ neurons) resulted in a longer gut transit time [*Edf1* gRNAs: 210.0 ± 34.9 min versus nontargeting (NT) gRNAs: 152.4 ± 13.1 min] and lower fecal water content (*Edf1* gRNAs: 57.3% ± 2.9% versus NT gRNAs: 65.4% ± 2.2%) (Fig. 6J), which are both consistent with slower intestinal transit. By contrast, deletion of *Mitf* (predicted as a regulator of Dcc^+ neurons) and *Asic4* (a regulator of $Epha8^+$ neurons) led to shorter gut transit times (*Mitf* gRNAs: 120.4 ± 8.7 min; *Asic4* gRNAs: 112.4 ± 9.0 min) and higher fecal water content (*Mitf* gRNAs: 68.3% ± 2.2%; *Asic4* gRNAs: 70.1% ± 4.2%) (Fig. 6J). To link these phenotypes to PIMN composition, we quantified $BGLAP^+$ and DCC+ neurons and observed that *Edf1* sgRNA–injected animals exhibited an increase in $BGLAP^+$ PIMN frequencies (Fig. 6K, left, and fig. S13), whereas *Mitf* sgRNA–injected animals demonstrated an increase in DCC^+ PIMN (Fig. 6L, left, and fig. S13). This effect was restricted to tdTomato-$KASH^+$ cells (Fig. 6, K and L, right), suggesting a direct link to the presence of the guide and deletion of *Edf1* or *Mitf*, whereas tdTomato-$KASH^-$ cells frequencies were unaffected compared with those of nontargeting control guides (Fig. 6, K and L, middle). These experiments validate the key regulators identified in our screen and directly link these candidate genes with neuronal state proportions and gastrointestinal transit time. Although these experiments were performed at steady state, the genes highlighted could potentially be targeted to modulate transit time in the context of gnotobiotic animals or allergy, which are characterized by a shift in the proportion of these inhibitory motor neurons.

Our in vivo ENS Perturb-seq enabled the rapid assessment of dozens of candidate genes in a pooled context and the functional validation of several transcription factors as potential regulators of PIMN differentiation along the continuum of cell states identified in gnotobiotic and type 2 immunity models. This work underscores the value of targeting ENS plasticity to better understand its role in disease.

Discussion

Here, we presented a single-cell–resolved map of enteric neuron adaptation to microbial colonization and type 2 inflammation. These two distinct perturbations modulated the ENS in shared ways, largely through transcriptional shifts in motor neurons and *Nmu*-hi sensory neurons. The latter were acutely responsive to both short-term microbial colonization and type 2 inflammation signals, which is perhaps expected given their extensive interactions with the immune system (*16*, *17*, *58*), and may be mediated by their ability to respond to type 2 cytokines and leukotrienes. In vivo perturbations further identified *Edf1* and *Mitf* as regulators of PIMN–cell state transitions and, concurrently, of gastrointestinal transit time. These results suggest that modulating PIMN states can directly alter intestinal physiology and highlight AAV-Perturb-seq as a scalable platform to interrogate the ENS. Unlike most other neuronal systems, the ENS is capable of neurogenesis and differentiation throughout adult life, potentially through the retention of neurogenic potential in enteric glial cells (*59*). As our analyses identified *Bglap*-hi PIMN as a potential early differentiation state, it will be important to determine whether the changes in motor neuron states we observed is directly correlated with modulation of neurogenesis across tissues and perturbations.

Unlike motor neurons, sensory neurons demonstrated strikingly distinct transcriptomic—and likely functional—profiles. We identified a *Grp*-hi sensory population enriched in the GF colon, especially under GF conditions. Our findings suggest that this population may interact with ICC through GRPR to regulate peristalsis. *Mc4r*-hi neurons also showed a distinct transcriptomic profile and expressed *Cck*. In a recent study, DREADD-based activation of enteric Cck^+ neurons showed reduced feeding behavior and fecal events (*14*). Our data shows that *Cck* is expressed in both *Mc4r*-hi and *Piezo2*-hi sensory neurons, but *MC4R*'s association with obesity and feeding behavior suggests that these neurons may be part of a network involving intestinal sensing and appetite regulation (*60*, *61*).

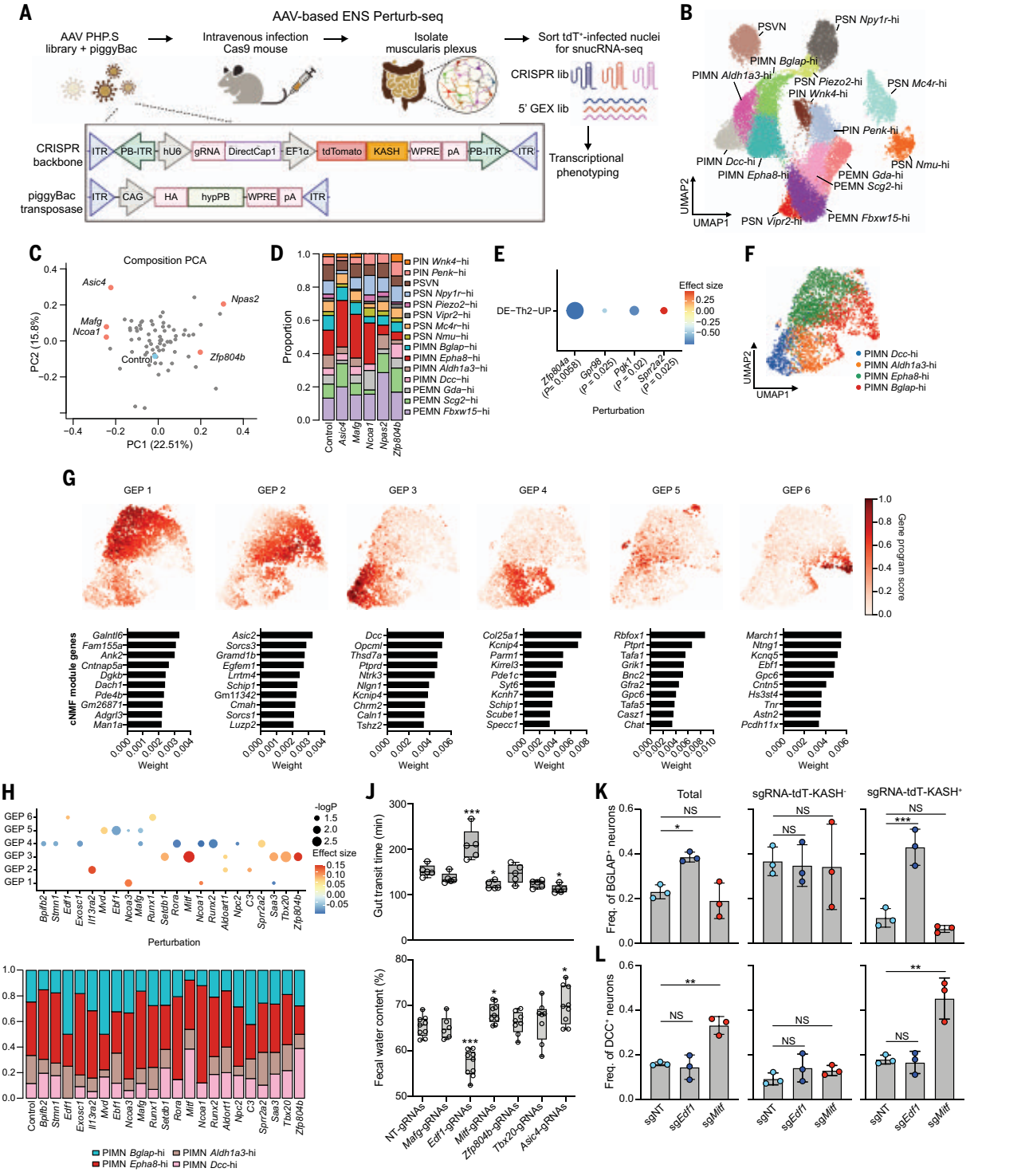


Fig. 6. In vivo Perturb-seq reveals key regulators of PIMN differentiation trajectory and neuronal programs. (**A**) Schematic diagram showing the experimental pipeline of AAV-based ENS Perturb-seq. The CRISPR backbone structure and piggyBac transposon elements are shown. CAG, CAG promoter [cytomegalovirus (CMV) enhancer fused to chicken β-actin promoter]; DirectCap1, DirectCapture sequence 1; HA, hemagglutinin A; hu6, human U6 promoter; hypPB, hyperactive piggyBac; ITR, inverted terminal repeat; KASH, Klarsicht/ANC-1/SYNE homology; pA, polyadenylation; PB-ITR, piggyBac ITR; WPRE, Woodchuck hepatitis virus posttranscriptional regulatory element. (**B**) UMAP embedding of enteric neuron nuclei transduced with AAV-PHP.S and isolated from the mouse muscularis propria. (**C**) PCA plot based on cell type composition. Each dot represents the composition of neurons assigned to guides from one target gene screened in the in vivo Perturb-seq library. Targets whose perturbation induced a significant change in cell type composition of enteric neurons are shown as labeled red dots. Nontargeting control is highlighted in blue. (**D**) Cell type proportion plots for target perturbations that were significant in (C), along with the nontargeting control. (**E**) Dot plot showing the effect sizes on the type 2 immunity–induced signature in enteric neurons after knocking out a target gene (*x* axis). Only associations with $P < 0.05$ were considered significant, with *P* values shown on the *x*-axis labels. (**F**) Recomputed UMAP embedding of only PIMNs with a single gRNA perturbation. Clustered PIMNs are labeled according to the indicated colors. (**G**) (Top) Gene program scores for the six gene expression programs (GEPs) identified by cNMF analysis are represented on the PIMN UMAP embedding shown in (F). (Bottom) Weights of the top 10 genes contributing to the highest loadings of the

same six GEPs. (**H**) Dot plots summarizing the association of the effect of knocking out a target gene (*x* axis) and its effect on the PIMN GEP scores. Only associations with $P < 0.05$ are considered significant. (**I**) Plot of PIMN cell type proportions for the neurons expressing guides specific for the same target genes as shown in (H) or the corresponding control (leftmost bar). (**J**) Measurement of gut transit time (top, $n = 5$ mice per group; *P* values: *Edf1* = 0.0003; *Mitf* =0.0243; *Asic4* = 0.0188) and fecal water content (bottom) expressed as a percentage of total stool weight ($n = 6$ for sg*Mafg*; 8 for sg*Zfp804b* and sg*Tbx20*; 9 for NT, sg*Edf1*, sg*Mitf*, and sg*Asic4*; *P* values: *Edf1* = <0.0001; *Mitf* = 0.0444; *Asic4* = 0.0290) in animals injected with AAVs carrying gRNAs targeting the indicated genes. (**K** and **L**) Quantification of the frequencies of ileal (K) BGLAP$^+$NOS1$^+$ or (L) DCC$^+$NOS1$^+$ neurons (number/mm^2) in total neurons, non–sgRNA-transduced neurons (sgRNA-tdT-KASH$^-$), and sgRNA-transduced neurons (sgRNA tdT-KASH$^+$) ($n = 3$ Cas9 mice) shown in fig. S13, across control (sgNT) or the indicated gene deletions (sg*Edf1* and sg*Mitf*). *P* values [versus sgNT, in (K)]: Total, sg*Edf1* = 0.0225, sg*Mitf* = 0.6237; tdT$^-$, sg*Edf1* = 0.9821, sg*Mitf* = 0.9698; tdT$^+$, sg*Edf1* = 0.0008, sg*Mitf* = 0.5240. *P* values [versus sgNT, in (L)]: Total, sg*Edf1* = 0.8861, sg*Mitf* = 0.0041; tdT$^-$, sg*Edf1*=0.3774, sg*Mitf* = 0.5412; tdT$^+$, sg*Edf1* = 0.9637, sg*Mitf* = 0.0038.

We expect that combining single-cell profiling and in vivo phenotypic screens at scale will enable the field to rapidly determine and prioritize regulators of ENS function, with applications in multiple immune and functional disorders of intestinal physiology.

Materials and methods

Experimental procedures

Animal husbandry: Animal experiments in this study were approved and carried out following the protocols (0055-05-15, 0122-10-16 and 0216-07-18) provided by the IACUC at the Broad Institute of MIT and Harvard, and protocol 2003N000158 provided by the IACUC at Massachusetts General Hospital. IACUC uses the NIH Guide for the Care and Use of Laboratory Animals, which is based on the US Government Principles for Utilization and Care of Vertebrate Animals Used in Testing, Research, and Training. Adult wild-type C57BL/6, *Uchl1-H2B-mCherry* reporter Tg(Uchl1-HIST2H2BE/mCherry/EGFP*)FSout/J (016981,Jackson laboratory, ME, USA), Cas9-GFP B6J.129(Cg)-Gt(ROSA)26Sortm1.1(CAG-cas9*,-EGFP)Fezh/J (026179, Jackson laboratory, ME, USA) mice were maintained in our facilities. *Uchl1-H2B-mCherry* reporter mice were backcrossed for 10 generations on the C57BL/6 background, and the line was rederived to germ-free by aseptic cesarean delivery in our facility. Pups were fostered by wild-type germ-free mothers and used to establish a reporter line maintained in gnotobiotic conditions. For all experiments, 7-to-12-week-old female or male mice were randomly assigned to treatment groups after age and sex matching (table S1).

Microbial cultures and GF animal colonization: All *B. theta* (DSM 2079 type strain, 29148, ATCC, VA, USA) and ASF strains (ASF356 *Clostridium sp.*, ASF360 *Lactobacillus intestinalis*, ASF361 *Lactobacillus murinus*, ASF457 *Mucispirillum schaedleri*, ASF492 *Eubacterium plexicaudatum*, ASF500 *Pseudoflavonifractor sp.*, ASF502 *Clostridium sp.*, ASF519 *Parabacteroides goldsteinii*) were purchased from Iowa State University, USA and grown on Columbia agar base medium (R452954, Thermo Fisher Scientific, MA, USA), yeast extract (211929, Thermo Fisher Scientific, MA, USA), vitamin K-hemin (10 x 10 ml, Fisher, B12354, Fisher Scientific, NH, USA), cysteine (Fisher, AAJ6374514, Fisher Scientific, NH, USA), 5% newborn calf serum (16010167, Thermo Fisher Scientific, MA, USA), 5% horse serum (26050088, Thermo Fisher Scientific, MA, USA), and 5% sheep serum (S2263, Sigma-Aldrich, MO, USA) in a forced air incubator at 37°C inside an anaerobic chamber. Of note, ASF457 and ASF492 are slow replicating and oxygen-sensitive and bacterial growth on fresh media may require 48 to 72 hours. Six-to-eight-week-old GF WT mice (C57BL/6N, Taconic Biosciences, NY, USA) or re-derived GF *Uchl1-H2B-mCherry* reporter mice (in-house, Broad Institute of MIT and Harvard) were orally gavaged with a single-dose of 100 µl 10^9 CFU/ml bacteria or 200 µl SPF feces (25 mg feces homogenized in PBS) and analyzed 4 weeks after colonization. Successful colonization was confirmed by villi size, cecum size, IgA$^+$ plasma cell numbers and T cell immune profiling on mesenteric lymph nodes, and PCR on feces stool samples. The relative proportion of bacterial vs. host RNA in each preparation was defined by quantitative reverse transcription polymerase chain reaction (qRT-PCR), using primers directed at *B. theta* 16S rRNA (forward, GGTAGTCCACACAGTAAACGATGAA; reverse, CCCGTCAATTCCTTTGAGTTTC), ASF356 (forward, AAAATAATTAGGAGCTTGCTTTTAA; reverse, TTAGAAGATGCCTCCTAAGAACC), ASF360 (forward, GGTGATGACGCTGGGAAC; reverse, AAGCAATAGCCATGCAGC), ASF361 (forward, GAACGAAACTTCTTTATCACC; reverse, TAGCATAGCCACCTTTTACA), ASF457 (forward, TCTCTTCGGGGATGATTAAAC; reverse, AACTTTTCCTATATAAACATGCAC), ASF492 (forward, AATTCCTTCGGGGAGGAAGC; reverse, TAAAACCATGCGGTTTTAAAAAC), ASF500 (forward, ACGGAGGACCCCTGAAGG; reverse, AGCGATAAATCTTTGATGTCC), ASF502 (forward, GAGCGAAGCACTTTTTTAGAAC; reverse, TTACACCACCTCAGTTTTTACC), ASF519 (forward, GCAGCACGATGTAGCAATACA; reverse, TTAACAAATATTTCCATGTGGAAC), and mouse 18S rRNA (forward, CATTCGAACGTCTGCCCTATC; reverse, CCTGTGCCTTCCTTGGA).

Oral food allergy model: Six-to-eight-week-old C57BL6/J mice were sensitized with a mixture of equal parts OVA solution (A5378, Sigma-Aldrich, MO, USA) and Alhydrogel 2% (vac-alu-50, InvivoGen, CA, USA) on days 0, 7, and 14. OVA/Alhydrogel mixture was made freshly on the day of injection and 100 µl of the mixture (50 µg OVA + 1%Alhydrogel) was injected into the scruff of animals. Blood was drawn from the facial vein at 12-24 hours before challenge, and 90 min after challenge. For the challenge, mice were orally gavaged with 50 mg OVA (A5503, Sigma-Aldrich, MO, USA) in 300 µl PBS every other day from Day 28 to Day 38, or until OVA-specific IgE and elevated total IgE were detected by serum enzyme-linked immunosorbent assay (ELISA). For IL-4/IL-13 signaling blockade, 50 µg anti-IL-4Rα (MAB530, R&D Systems, Minneapolis, MN, USA) was intraperitoneally injected once per week from day 0 to day 28. Beginning on day 28, 50 µg anti-IL-4Rα and OVA challenge (retro-orbital vein injection of 100 µg OVA) were administered together every other day until day 36.

Systemic food allergy model: Six-to-eight-week-old C57BL6/J mice were sensitized by oral gavage of 5 mg OVA and 10 µg cholera toxin on days 0 and 7, followed by retro-orbital vein injection of 100 µg OVA 14 days after first sensitization.

Fecal water content measurements: We quantified fecal water contents by collecting fresh fecal pellets and weighing that at time of collection and after drying in a lab oven at 60° to 65°C for 24 to 48 hours. Fecal water percentage was calculated as (fresh weight - dried weight)/fresh weight × 100.

Carmine-red dye for measuring gastrointestinal transit times: Gastrointestinal transit times were measured using nonabsorbable carmine-red dye (Sigma-Aldrich, #C1002; St. Louis, MO), prepared as a 6% (w/v) solution in 0.5% methylcellulose (Sigma-Aldrich, #M0512; St. Louis, MO), and autoclaved before being introduced into gnotobiotic isolators. Mice were maintained on a strict 12-hour light cycle (lights on between 06:00 and 18:00) and gavaged with 150 µl of the carmine solution between 08:00 and 08:30. The mice were not fasted prior to the experiment. Feces were collected every 30 min for up to 8 hours postgavage, and each sample was streaked on a sterile white napkin to detect the presence of red carmine dye. The time from gavage to the first appearance of carmine in the feces was recorded as the total gastrointestinal transit time for each animal.

***H. polygyrus* infection:** *Uchl1*-reporter or C57BL6/J WT mice were infected with 200 *H. polygyrus* L3 larvae in 200 µl of distilled water by oral gavage on day 0, maintained under specific-pathogen-free (SPF) conditions at Massachusetts General Hospital. Mice were euthanized 14 days after *H. polygyrus* infection. Note that on day 1 to 7, worms invade the duodenum wall and encyst in the subserosa; on day 8 to 14, adult worms return to duodenum lumen for mating and egg production. At necropsy, the intestine was removed from *Uchl1*-reporter or C57BL6/J WT mice, and worms were removed and counted. Notably, *Uchl1*-reporter mice showed comparable worm burden on day 14 (fig. S1I) to WT mice.

Swiss rolling of gut tissue segments: The small intestine (duodenum, jejunum and ileum) and colon tissues were isolated from mice. Tissues were cut into 6-cm segments and opened longitudinally, gently washed with ice-cold PBS, and flattened with a microhematocrit capillary tube (15401-560, VWR, PA, USA) on an ice-cold metal block with proximal side on the top, distal side at the bottom, and smooth serosa side up. A damp tissue paper was laid on top of the tissues and flipped around with tissue luminal side up. The distal end of tissue (bottom) was rolled onto a microhematocrit capillary tube (Z114952, Sigma-Aldrich, MO, USA) and rolled upward to the proximal end of the tissue (top). The tissue roll with a capillary tube was turned upright and put into a plastic tissue mold on a freezing cold metal block immersed in liquid nitrogen. Optimal cutting temperature compound (Tissue-tek O.C.T. compound, 4583, Sakura Finetek, CA, USA) was slowly filled into the mold to cover the whole tissue roll. The capillary tube was pulled out when the OCT started to freeze.

Isolation of submucosal and myenteric plexuses: Isolation of submucosal and myenteric plexuses were performed as described previously (*62*). Briefly, gut tissues were cut into 6-cm segments and opened longitudinally, gently washed with ice-cold PBS, and flattened with a microhematocrit capillary tube on an ice-cold metal block with the smooth serosal side up. Connective tissue and vessel attachments were gently removed. The beginning of the tissue segment was pressed down with a small damp paint brush, and a pair of curved tweezers was used to gently make a dent in the tissue. The tissue from the indentation was scraped carefully with the curved tweezers until the layers separated. The submucosal and myenteric plexuses were scraped away until the end of the length of tissue was reached.

Immunofluorescence (IF) staining: For whole mount IF, following intestine dissection and washing, whole intestine tissues or submucosal and myenteric plexuses were isolated and pinned down on a plate coated with Sylgard, followed by overnight fixation in 4% paraformaldehyde (PFA)/PBS at 4°C. After washing in PBS, samples were then permeabilized in the permeabilization buffer (PBS with 0.5% Triton X-100, 0.05% Tween-20, 4 µg/µl heparin) for 2 hours at room temperature (RT) with gentle agitation. Samples were then blocked for 2 hours in a permeabilization buffer containing 5% normal goat serum (5425, Cell signaling technology, MA, USA) for 2 hours at RT with gentle agitation. Primary antibodies were added to the antibody diluent (8112, Cell signaling technology, MA, USA) at appropriate concentrations and incubated for 2 days at 4°C. After primary incubation samples were washed in the permeabilization buffer, followed by incubation with the secondary antibody in antibody diluent at appropriate concentrations for 2 hours at RT. Samples were again washed in a permeabilization buffer, and then mounted with fluoroshield mounting medium with DAPI (ab104139, Abcam, UK) and coverslips. Slides were kept in the dark at 4°C until they were imaged. For cross-sectional IF staining of Swiss-roll cryosections, or for staining of neurons after calcium imaging, time for fixation, primary and secondary antibody incubation was shortened to 1 hour.

The following antibodies were used in this study: Goat anti-Mouse IgE-HRP (SouthernBiotech Cat# 1110-05, RRID:AB_2794604), anti-PTGIR (Alomone Labs Cat# APR-068, RRID:AB_2756763), anti-MC4R (Bioss Cat# BS-11417R), anti-GPR158 (Sigma-Aldrich Cat# SAB4502509, RRID:AB_10762838), anti-Osteocalcin/BGLAP (Proteintech Cat# 23418-1-AP, RRID:AB_2879275), anti-GRP (ImmunoStar Cat# 20073, RRID:AB_572221), anti-TAC1 (Atlas Antibodies Cat# HPA014429, RRID:AB_1857722), mouse anti-TUBB3 (BioLegend Cat# 801202, RRID:AB_2313773), anti-IL-36/IL-1F6 (R and D Systems Cat# AF2297, RRID:AB_355216); anti-MAP2 (ab32454, RRID:AB_776174), rabbit anti-NMU (ab229635, RRID: AB_3699305), anti-VIP (ab8556, RRID:AB_306628), chicken anti-TUBB3 (ab41489, RRID:AB_727049), anti-EGFR (ab52894, RRID:AB_869579), anti-GAL (ab99452, RRID:AB_10672709), anti-HuC/D (ab184267, RRID:AB_2864321) and anti-CHAT (ab18736, RRID:AB_2244867) were purchased from Abcam; mouse anti-NMU (sc-398600, RRID: AB_3699304), anti-ANO1 (sc-377115) and anti-NOS1 (sc-5302, RRID:AB_626757) were purchased from Santa Cruz Biotechnology, TX, USA; anti-TAU (13-6400, RRID:AB_2533025), FITC-anti-IgA(11-4204-81, RRID:AB_465220), anti-CXCL12 (MA5-23759, RRID:AB_2608711), anti-IGFBP4 (PA5-25925, RRID:AB_2543425), anti-GDA (PA5-30322, RRID:AB_2547796), anti-DCC (PA5-115098, RRID:AB_2899734), anti-LGR5 (MA5-25644, RRID:AB_2723318), anti-ALDH1A3 (PA5-29188, RRID:AB_2546664), anti-SST (MA5-16987, RRID:AB_2538460), anti-PENK (PA1-30928, RRID:AB_1958811), anti-SCTR (PA5-102701, RRID:AB_2852095), anti-CCK (PA5-103116, RRID:AB_2852486), anti-PTCH2 (PA1-46223, RRID:AB_2174052), anti-MOXD1 (BS-17733R), anti-GRPR (PA5-33834, RRID:AB_2551204), anti-CysLTR2 (PA5-75328, RRID:AB_2719056), and Alexa Fluor 647-anti-mCherry (M11241, RRID:AB_2536615) were purchased from Thermo Fisher Scientific, MA, USA. Antibodies were validated based on their specificity (low or no background, with only a fraction of cells having positive staining signals), cellular location of the staining, neurotransmitters' and neuropeptides' colocalization with neuronal markers, manufacturer reports, and use in prior publications.

For quantification of enteric neurons by confocal imaging, a minimum of 5 to 10 images per biological replicate were randomly acquired across a piece of whole mount submucosal and myenteric plexuses, and at least three replicates were used for each condition. Images were opened in Fiji (ImageJ, NIH, USA), and the number of cells in a given field was manually counted. This number was then multiplied by a factor of 2.95 (20x objective) to calculate the number of counted neurons per square millimeter (mm^2). The average of five (or more) images were then calculated and plotted. Every point on a given graph corresponds to a single animal. Data were plotted with GraphPad Prism10.

Nucleus extractions for single-nucleus RNA-sequencing: Nucleus extraction protocols were performed on RNase-free benchtop and buffers were prepared freshly on the day performing nucleus isolation.

Buffer compositions for each protocol are listed: CST buffer: 0.49% (w/v) CHAPS (10810118001, Sigma-Aldrich, MO, USA) in ST buffer (146 mM NaCl, 1mM CaCl2, 21 mM MgCl2,10 mM Tris pH 8.0), 20 mg/ml BSA (B9200, New England Biolabs, MA, USA), 1x protease inhibitor cocktail (11873580001, Roche, Switzerland), 0.2 to 1 U/µl Protector RNase inhibitor (3335402001, Roche, Switzerland). Nucleus isolation buffer (NIB): 250 mM sucrose (S0389, Sigma-Aldrich, MO, USA), 25 mM KCl, 5 mM MgCl2 and 10 mM Tris (pH 8.0) in nuclease-free H_2O. Can be stored for up to 3 months at 4°C. Homogenization buffer (HB): 1 mM DTT (R0861, Thermo Fisher Scientific, MA, USA), 1x protease inhibitor cocktail, 0.1% Triton-X (vol/vol) and 500 ng/ml DAPI in NIB.

For chopping extraction for SMART-seq2, flash-frozen tissues were disaggregated in 1 ml of CST buffer with quick chopping by Tungsten Carbide Straight 11.5-cm Fine Scissors (14558-11, Fine Science Tools, CA, USA) for 10 min on ice in a six-well plate. Large debris was removed with a 40-µm cell strainer (21008-949, VWR, PA, USA). An additional 1 ml of CST buffer was used to wash the well and strainer. Three milliliters of ST buffer was used to wash the strainer to extract nuclei. Nuclei were then pelleted at 500 g for 5 min at 4°C. Supernatant

was discarded and the nuclei pellet was resuspended in 100 to 500 µl of ST buffer before proceeding to fluorescence-activated cell sorting (FACS).

For dounce homogenization (*63*), tissues were transferred into a 1.5-ml tube containing 500 µl of HB with DAPI added (1µg/ml) and minced with surgical scissors until they were small enough to collect with a P1000 pipette. The DAPI was added to label DNA and allow identification of intact nuclei by FACS. Tissue contents of the tube were then transferred to a dounce homogenizer containing 1 ml of ice-cold HB. To ensure that all tissue was collected, the 1.5-ml tube was rinsed with an additional HB (1 to 1.5 ml) and these contents were also transferred to the homogenizer. Tissues were manually homogenized using the homogenizer pestle (10 to 15 strokes) and visually inspected to ensure no large pieces remained (the number of strokes needed for successful homogenization varies according to each intestinal segment; as colonic tissues require a greater number of strokes than SI. Excessive homogenization can result in sheared nuclei, while insufficient homogenization leads to less efficient isolation of nuclei). Homogenized tissues were then vortexed for 20 s. The homogenate (containing the neuronal nuclei and cellular debris) was filtered through a 70-µm pore size cell strainer (21008-952, VWR, PA, USA) into a new sterile 50-ml conical tube. The tissues were mashed on the strainer using the rubber side of a 1-ml syringe plug, and 500 µl to 1 ml of HB was added to completely collect the nuclear samples. Samples were kept on ice throughout this process. The filtered homogenate was then centrifuged at 1000 g for 10 min at 4°C and the supernatant carefully discarded. Nuclei were resuspended in a solution of 1x PBS, 1% BSA (130-091-376, Miltenyi Biotec, Germany), and 1 U/µl RNase inhibitor.

SMART-seq2: RNA lysate clean-up was performed using RNAClean XP beads (A63987, Beckman Coulter, CA, USA) followed by reverse transcription with Maxima H minus Reverse Transcriptase (EP0743, Thermo Fisher Scientific, MA, USA) and whole-transcription amplification (WTA) with KAPA HotStart HIFI 2 × ReadyMix (07958927001, Roche, Switzerland) for 21 cycles. WTA products were purified with Ampure XP beads (A63881, Beckman Coulter, CA, USA), quantified with Qubit dsDNA HS Assay Kit (Q32854, Thermo Fisher Scientific, MA, USA) and assessed with D5000 ScreenTape on a 2200 Agilent TapeStation system (5067-5588, 5067-5589, Agilent, CA, USA). Diluted WTA products underwent tagmentation and indexing using the Nextera XT DNA Library Preparation Kit (FC-131-1096, Illumina, CA, USA), followed by size selection with Ampure XP beads. Size and concentration were then assessed with D1000 screentape (5067-5582, 5067-5583, Agilent, CA, USA) on 2200 Agilent TapeStation system. Libraries from four 96-well plates were pooled together and sequenced with the Illumina NextSeq 500/550 75-cycle high output kit (20024906, Illumina, CA, USA) on a NextSeq 500/550 system.

Large intestine cell isolation for single-cell RNA-seq from GF and SPF mice: To obtain all cell types from the intestine for single-cell RNA-seq, we collected intestine segments from GF and SPF female C57BL6/J mouse pairs in ice-cold 2% FBS/PBS. After removing fat tissue, segments were cut open longitudinally and washed with ice-cold PBS. Epithelial cells were stripped with a stripping buffer (5 mM EDTA, 1 mM DTT, 2.5 mM HEPES, and 5% FBS in HBSS) for 30 min at 37°C with stirring. Tissues were collected and rinsed with ice-cold 2% FBS/PBS, followed by enzymatic digestion with 5 ml digestion buffer (0.2 mg/ml collagenase IV, 0.2 mg/ml dispase II (17105041,Thermo Fisher Scientific), 0.1 mg/ml DNase I in DMEM/F12) at 40°C for 10 min. Digested tissues were allowed to rest for 5 min, and the supernatants were collected into ice-cold complete media (CO_2 independent media, 18045088, Thermo Fisher Scientific; with 2% FBS) with 10 mM EDTA and kept on ice. This process was repeated for approximately 45 min with gentle vortex until no tissue chunks remained. Cells were pooled together and filtered through a 70-µm cell strainer, centrifuged at 500 g at 4°C for 10 min. Cell pellets were resuspended in complete media (CO_2 independent media with 2% FBS) with 10 mM EDTA. At this stage, 150 µl of the suspension was used for immediate loading ("bulk"), while the remaining 850 µl were enriched for stromal cells ("stroma"). Enrichment of stromal cells was performed by staining the cells with PE-Cy7 anti-Ter-119 (BioLegend Cat# 116222, RRID:AB_2281408, 1:100 dilution), PE-Cy7 anti-CD45 (BioLegend Cat# 103114, RRID:AB_312979, 1:100), FITC anti-EpCAM (Santa Cruz Biotechnology Cat# sc-53532, RRID:AB_2277892, 1:50) in the presence of FcR blocking reagent (Miltenyi Biotec Cat# 130-092-575, RRID:AB_2892833, 1:50) for 20 min on ice. NucBlue (Invitrogen, R37605, one drop per sample) was added before sorting to exclude dead cells and live, marker negative cells were sorted on a MoFlo Astrios EQ (Beckman Coulter) as the stromal fraction.

For both the bulk and stromal fractions, cell number was counted and the resuspension was diluted to a final concentration of 20,000 cells total in 75 µl before loading into a Chromium controller (10x Genomics, CA, USA), followed by library construction using the Chromium Single Cell 3′ v2 kit (PN-120237, 10x Genomics, CA, USA).

AAV-PHP.S packaging and peripheral neuron transduction: AAV-PHP.S viruses were produced and purified according to published protocols (*64*). Briefly, AAV-PB-tdT-KASH-DC-perturb-seq (constructed in this paper) or AAV-HA-hypPB (*65*) (containing piggyBac transposase; gifted from Xin Jin, Scripps Research), pUCmini-iCAP-PHP.S (103006, Addgene, MA, USA), pHelper (81070, Addgene, MA, USA) plasmid DNAs were transfected into AAVpro 293T cells (632273, Takara Bio, Japan) at ratio = 1:4:2 in a total amount of 40 ug per 150-mm dish with polyethylenimine (PEI; molecular weight 25,000; 23966-1, Polysciences, PA, USA) stock solution (7.5 mM; DNA:PEI=27.5:1). Supernatants were collected at 72 hours and 120 hours posttransfection, and viral particles in supernatants were precipitated with ⅓ volume of 40% polyethylene glycol (PEG) on ice for 2 hours, followed by centrifugation at 4000 g, 4°C for 30 min. PEG pellets and cells from dishes were combined and lysed with 100 U/ml salt-active nuclease (HL-SAN; 25 U/µl; 70910-202, ArcticZymes, Norway) in SAN digestion buffer (500 mM NaCl, 40 mM Tris base, and 10 mM MgCl2, pH 10.0) at 37°C water bath for 1 hour and 30 min. The lysate was centrifuged at 2000 g, RT for 10 min. Supernatants were loaded onto a iodixanol density gradient (60, 40, 25, and 15%) and centrifuged at max speed (50,000 rpm) for 2 hours and 30 min at 18°C. AAV particles were pulled out from the yellowish layer between the 60 and 40% gradients, filtered through a 0.22-µm filter and a 100-kDa amicon filter device (UFC9100, Sigma-Aldrich, MO, USA), and washed twice with 1x PBS+ 0.001% (v/v) pluronic F-68 (24040-032, ThermoFisher Scientific, MA, USA) to retain 300 to 500 µl of solution in the top chamber, containing AAVs ready for injection.

An aliquot was used for AAV titration as follows: AAVs were digested with 50 U/ml DNase I in DNase digestion buffer (2 mM $CaCl_2$, 10 mM Tris-HCl, and 10 mM $MgCl_2$) at 37°C water bath for 1 hour, followed by EDTA incubation at 70°C for 10 min, and incubated with 100 µg/ml proteinase K in proteinase K solution (1 M NaCl and 1% (w/v) N-lauroylsarcosine sodium salt) at 50°C for 2 hours. Titers determined by qPCR using primers targeting WPRE (forward, 5′-GGCTGTTG GGCACTGACAAT-3′; reverse, 5′-CCGAAGGGACGTAGCAGAAG-3′), compared to a linearized positive control plasmid.

A single dose of AAVs containing gRNA pools or gRNAs targeting a single gene (table S9) was intravenously injected into Cas9 mice at 2×10^{12} vg/mouse. Transduced mice were euthanized at 3 to 4 weeks postinjection.

Adult neuronal culture: Adult enteric neurons directly isolated from mice were cultured. For adult culture, buffers and culture media were first prepared as follows: Enteric neuronal culture medium: 2% FBS (A3840001, Thermo Fisher Scientific, MA, USA), 5% Chicken embryo extract (IC092850145, MP biomedicals), 1x Penicillin-Streptomycin-Amphotericin (516104-M, Sigma Aldrich, MO, USA), 1x Gentamicin, 1x N2

supplement, 1x B27 supplement in DMEM/F12. Collagenase digestion medium: 1 mg/ml collagenase IV (LS004188, Worthington Biochemical, NJ, USA), 0.5 mM CaCl2, 10 mM HEPES in HBSS. Digestion neutralizing medium: 400 U DNase I, 1 mg/ml BSA in DMEM/F12. Adult wild-type C57BL/6J mice (7 to 12 weeks old, male and female) were euthanized and intestinal longitudinal muscle/submucosal plexus/myenteric plexus were isolated and placed in a holding buffer. Tissues were digested in prewarmed collagenase digestion medium and incubated in 37°C water bath shaker (40 to 60 rpm) for 15 min (tissues were inverted several times every 5 min). Dissociated tissues were spun down at 200 g, 4°C for 5 min. The pellet was transferred to a 10-cm petri dish containing 10 ml HBSS and shaken for 10 min at RT. HBSS shaking was repeated and cells were spun down at 200 g 4°C for 5 min. Cells were then trypsin digested in a 37°C water bath shaker (40 to 60 rpm) for 10 min. Digestion was neutralized with digestion neutralizing medium and cells were triturated vigorously with a P1000 pipette. Cells were filtered, washed with HBSS, and spun down at 200 g, 4°C for 5 min. Cells were resuspended in enteric neuronal culture medium and seeded into 24-well plates (40,000 cells per well). The medium was changed every other day. Neurons were visible by day 2 to 3 and extensive around day 5.

Ca^{2+} imaging: The Fluo-4 Direct calcium assay kit (F10471, Thermo Fisher Scientific, MA, USA) was used to detect neuronal cell activation and Ca^{2+} influx, following the manufacturer's instructions. 250 mM stock solution of probenecid was prepared by adding 1 ml of Fluo-4 Direct calcium assay buffer to each 77-mg vial of water-soluble probenecid. 2X Fluo-4 Direct calcium reagent (with a final probenecid concentration of 5 mM) was made by mixing 10 ml Fluo-4 Direct calcium assay buffer and 200 µl 250 mM probenecid stock solution. An equal volume of 2X Fluo-4 Direct calcium reagent loading solution was added directly to wells containing cells in culture medium. Plates were incubated at 37°C for 30 to 60 min. Diluted ligands at the indicated working concentrations were added to the cell cultures during a time-lapse fluorescence recording, using instrument settings appropriate for excitation at 494 nm and emission at 516 nm. IL-13 (100 ng/ml; 575902, BioLegend, CA, USA) or leukotriene LTC_4 (100 nM; 20210, Cayman, MI, USA) treatment was used to determine the activation of $IL13RA1^+$ or $CysLTR2^+$ cells, respectively. KCl (50 mM) was used as a positive control to verify neuronal viability/activity. Following Ca^{2+} imaging, ligands were washed out and replaced with fresh culture media to allow cells to return to a resting state. Cells were then fixed and subjected to immunofluorescence staining using antibodies against CHAT, NOS1, and NMU to confirm their identity [see "IF staining" above]. IL-13 responsiveness, LTC_4 responsiveness and NMU positivity were quantified across the KCl-responsive cells in 12 total imaging areas, and numbers of cells positive for 1, 2, or 3 of these categories is presented as a Venn diagram.

Fluorescence-activated cell sorting (FACS) and flow cytometry analysis: Single-nucleus suspensions were obtained from flash-frozen tissues. Sorting and FACS analysis were performed on a MA900 multi-application cell sorter (Sony Biotechnology, CA, USA), and data were analyzed by FlowJo software. Nuclei population was gated using FSC/SSC and doublets were excluded by gating on FSC-H/FSC-A. $mCherry^+$ or $tdTomato^+$ and $DAPI^+$ populations were gated for enteric neurons. Sorting was performed on a Sony MA900 sorter in 96-well plate mode using ultra-pure, single-nucleus sorting, one nucleus per well for SMART-seq2, and 96-well plate mode, normal purity sorting for Perturb-seq.

ELISA for cytokines and IgE amounts: Serum and stool samples were collected at indicated time points after challenges. Cytokine concentrations were determined by ELISA with primary capture and secondary biotin-conjugated antibodies against mouse IL-4 (Cat# MM450DB, RRID:AB_223604 and Cat# MM450C, RRID:AB_223530), IL-5 (# 14-7052-85, RRID:AB_468421 and Cat# 13-7051-85, RRID:AB_466907), IL-13 (Cat# 14-7133-85, RRID:AB_763554 and Cat# 13-7135-85, RRID:AB_763556), TNF-α (Cat# 14-7423-85, RRID:AB_468492 and Cat# 13-7341-85, RRID:AB_466951) and IFN-γ (antibody pair from kit Cat# KMC4021), all from Thermo Fisher Scientific. In brief, 96-well plates were precoated with the capture antibody (1:500 in coating buffer) or 100 µl ovalbumin (10 µg/ml) at 4°C overnight. On the next day, the plate was washed with 0.1% Tween-20/PBS and blocked with 1%BSA/PBS diluent buffer for 2 hours at RT. Diluted supernatants and cytokine standards were then applied to the plate and incubated for 2 hours at RT. The plate was then washed and incubated with a biotin-conjugated detection antibody (1:1000 in 1% BSA/PBS diluent buffer) for 1 hour at RT. The plate was then washed and incubated again with poly-HRP streptavidin (1:5000 in the diluent buffer, Thermo Fisher Scientific) for 30 min. The plate was washed again and finally incubated with the tetramethylbenzidine (TMB) substrate solution (555214, BD Biosciences, NJ, USA). The reaction was stopped using 2 M H_2SO_4. The absorbance of each well was recorded by a Synergy 2 plate reader (7131000, BioTek, VT, USA) at 450 nm. The absorbance of the standard sample was used to construct the standard curve. For detection of total IgE and OVA-specific IgE, 96-well plates were precoated with IgE primary antibody (from Thermo Fisher Scientific Cat# 88-50460-22, RRID:AB_2574855) or OVA (20 µg) at 4°C overnight. On the next day, 4 µl of the serum was diluted in 100 µl of cold ELISA blocking buffer. One-hundred microliters of diluted samples were added to wells and incubated for 1 hour at RT or overnight in a cold room while shaking. Basal serum Ig amounts were measured using unvaccinated animal serum. Each plate was washed 4 times with PBS + 0.05% Tween-20. One-hundred microliters of blocking buffer containing goat anti-Mouse IgE-HRP (SouthernBiotech Cat# 1110-05, RRID:AB_2794604) was diluted at 1:2500 per well. After a further incubation for 1 hour at RT while shaking, the plate was washed another four times with 0.1% Tween-20/PBS and developed with TMB substrate solution.

Quantitative RT-PCR analysis: Total RNA was isolated cells or tissues using the RNeasy RNA isolation kit (74104, Qiagen, Netherlands) and reverse transcribed using SuperScript III reverse transcriptase (18080093, ThermoFisher Scientific, MA, USA) per the manufacturer's directions. Quantitative PCR was conducted on the Bio-Rad CFX384 using Power SYBR Green qPCR master mix (4367659, Applied Biosystems, MA, USA). mRNA amounts were normalized to *Gapdh*, and relative amounts were determined using the ΔΔCT method. Primers are listed in table S2. The following qPCR reaction mix was prepared: 2 µl cDNA, 5 µl PowerSYBR Green PCR Master Mix, 0.5 µl Primer mix (5 µM), 2.5 H_20 and run with the following thermal cycling conditions: hold at 95°C for 10 min, denature at 95°C for 15 s and anneal/extend at 60°C for 60 s for 40 cycles.

RNAscope: RNAscope Multiplex Fluorescent Reagent Kit V2 (323270, Advanced Cell Diagnostics, CA, USA) was used per manufacturer's recommendations for flash-frozen samples with the following alterations: All wash buffer times were increased to 5 min and, following the final HRP-Block step, slides were washed for 10 min with PBS containing DAPI followed by mounting with Fluoromount-G (0100-01, SouthernBiotech, AL, USA) and sealed. Probes used for smFISH were purchased from Advanced Cell Diagnostics: *Calca* (417961), *Calcb* (425511), *Cck* (402271), *Chat* (408731-C2), *Grp* (317861-C2), *Htr4* (408241), *Lgr5* (312171), *Nmu* (446831), *Nog* (467391), *Nos1* (437651-C3), *Piezo1* (500511), *Piezo2* (400191-C3), *Sst* (404631-C3), *Vip* (415961 and 415961-C2) were used. In the case of combining smFISH and IF staining, the following changes were made: after the final HRP-Block step, tissue sections were incubated with primary antibodies overnight at 4°C, washed three times in TBST for 5 min each, and then incubated with secondary antibodies for 30 min at RT. Slides were then washed three

times in TBST for 5 min each. Slides were mounted with a Fluoroshield mounting medium with DAPI and sealed.

In vivo Perturb-seq: AAV-PB-tdT-KASH-DC-perturb-seq vector was constructed based on px552 (Addgene, 60958), piggyBac right (3′) inverted repeat (CCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATC ATGTGTAAAATTGACGCATG) was inserted before U6 promoter, hSyn promoter was replaced with EF-1α promoter, GFP was replaced with tdTomato (tdT) sequence without stop codon fused with KASH, gRNA scaffold was replaced with gRNA scaffold from Perturb-seq backbone (*19*) (GTTTAAGAGCTAAGCTGGAAACAGCATAGCAAGTTTAAA TAA GGC TAG TCC GTT ATC AAC TTG AAA AAGTGGCA-CCGAGTCGGTGC), followed by a direct capture (DC) sequence 1 (*66*) GCTTTAAGGCCGGTCCTAGCAA. piggyBac left (5′) inverted repeat (CATGCGTCAATTTTACGCA GACTATCTTTCTAGGG) was inserted between hGH poly(A) signal and AAV2 ITR. sgRNA libraries listed in table S9 were cloned into AAV-PB-tdT-KASH-DC-perturb-seq at SapI site. Submucosal and myenteric plexus layers were isolated from AAV-PHP.S Perturb-seq transduced mice. Eight-thousand single nuclei were sorted based on tdTomato$^+$. Single nucleus RNA-sequencing and CRISPR-sgRNA detection were performed using Chromium Next GEM Single Cell 5′ Reagent Kits v2 (Dual Index) with Feature Barcode technology for CRISPR Screening (PN-1000263, PN-1000190, PN-1000451, 10x Genomics, CA, USA). Gene expression and CRISPR feature barcode libraries were constructed. CRISPR perturbation analysis declared the target gene of each gRNA from Read2 using CRISPR protospacer sequence and constant sequence TTCCAGCATAGCTCTTAAAC(BC). Cells containing non-targeting guides or empty guides were used as controls.

Data analysis and statistics

Statistical analyses: Multiple group comparisons in Fig. 3, C and I; Fig. 4, G and H; Fig. 5, C, F, and L; Fig. 6, J to L; and figs. S8C and S9A were performed with one-way ANOVA followed by Tukey's multiple comparison test. Two-group comparisons in Fig. 5, D, G, and M; and fig. S1, D, J, and K were performed with unpaired Student's *t* test. *P* values lower than 0.05 were considered significant. Statistical tests performed on RNA-seq datasets are discussed in the following sections.

SMART-seq2 FASTQ processing: FASTQ files were demultiplexed and aligned to the reference transcriptome mm10 (including the introns) and transcripts were quantified using RSEM, as previously described (*18*). RSEM count files from every experimental batch (total experiments = 47) were merged to generate a complete dataset of 16,222 cells × 23,420 genes for downstream analysis.

Generation of integrated SMART-seq2–based ENS atlas of the whole intestine: A previously published SMART-seq2 based atlas of the colonic ENS (*18*) was integrated with the count matrix generated in the current study and analyzed further using the scanpy package. Quality control measures were applied to exclude poor-quality cells; specifically, cells expressing <200 detected genes were filtered out. Further filtering was performed to remove data from individual mice with <10 total number of cells. Count matrices of each individual cell were read depth normalized using the logTP10K normalization procedure and report the number of transcripts per 10,000 transcripts in a cell. To minimize the effect of highly expressed genes, genes with >5% counts of the total count per cell were excluded in the computation of normalization factors. Subsequently, the top 2000 highly variable genes were identified for each study separately and then merged. After that, the normalized expression matrix was centered and the effect of total RNA count per cell was regressed out. Batch correction was then performed using the Harmony package (*67*), considering the following variables as covariates: Mouse + Sex + Experiment + Study + Genotype (Sox10/Uchl1). The top 40 principal components (PCs) were selected based on inspection of the elbow plot, followed by the construction of a *k*-nearest neighbor (k-NN) graph and the application of the Leiden clustering algorithm to cluster the cells, which were visualized by UMAP embedding algorithm (*68*). Post hoc analysis was performed on identified clusters to remove poor-quality cells – including clusters with low total RNA counts – or potential doublets expressing lineage markers of non-neuronal cells (*Epcam*, *Pdgfra*, *Des*, *Pecam1*, *S100b*, *Ptrpc*). An additional round of clustering and filtering was performed, for a final dataset of 7640 cells.

Annotation and characterization of SMART-seq2–based ENS clusters: We annotated neurons based on their expression of general subtype markers first, and subdivided each main class based on transcriptional heterogeneity and annotated based unique or highly abundant genes as follows: (i) 4 classes of putative excitatory motor neurons (PEMN:*Chat*$^+$*Tac1*$^+$): F-box protein (*Fbxw15*)-hi; phospholipid scramblase (*Plscr2*)-hi; secretogranin 2 (*Scg2*)-hi; and guanine deaminase (*Gda*)-hi cells. (ii) 5 classes of putative inhibitory motor neurons (PIMN: *Nos1*$^+$): netrin-1 receptor (*Dcc*)-hi; aldehyde dehydrogenase 1A3 (*Aldh1a3*)-hi; ephrin receptor (*Epha8*)-hi; Copine-7 (*Cpne7*)-hi; and osteocalcin (*Bglap*)-hi cells. (iii) 6 classes of putative sensory neurons (PSN; *Chat*$^+$CGRP$^+$): neuromedin U (*Nmu*)-hi; melanocortin 4 Receptor (*Mc4r*)-hi; gastrin releasing peptide (*Grp*)-hi; vasoactive intestinal peptide receptor 2 (*Vipr2*)-hi cells; piezo type mechanosensitive ion channel component 2 (*Piezo2*)-*hi*; and neuropeptide Y receptor Y1 (*Npy1r*)-hi cells. (iv) 2 classes of putative secretomotor/vasodilator neurons (PSVN; *Glp2r*$^+$): Galanin (*Gal*)-hi cells; and natriuretic peptide receptor A (*Npr1*)-hi *cells*. (v) 2 classes of putative interneurons (PIN; positive for *Penk, Egfr, Tac1*): proenkephalin (*Penk*)-hi; chloride sensor (*Wnk4*)-hi cells.

Processing and analysis of Germfree and SPF mice: Gene expression matrices for every sample were obtained by aligning the FASTQ sequence reads to the reference mm10 transcriptome using CellRanger 3.1.0 software (10x Genomics, CA, USA).

Generation of an integrated pan-intestine whole-tissue scRNA-seq atlas in mice: Previously published scRNA-seq studies profiling the small intestine, duodenum and ileum, or large intestine of mice were considered for integrative analysis. Specifically, the following datasets were downloaded from GEO: GSE152325, GSE163638, GSE190037, GSE198465, GSE201723, GSE213424, GSE221006, GSE232703, GSE243483, GSE92332, GSE124880, GSE235743, GSE106510. Study details have been provided in fig. S5 and table S4. In general, we prioritized previous studies studying type 2 inflammation conditions. However, since stromal cells were not previously studied in the context of type 2 inflammation, we included data from other conditions to generate a whole-tissue cell type atlas. We also included the scRNA-seq data generated for GF and SPF mice, as described above.

Count matrices were combined to generate an integrative dataset for further analysis using the scanpy package. Cells expressing <200 detected genes were filtered out as a first pass QC measure, followed by gene expression normalization to account for differences in sequencing depth across cells. In general, the downstream analysis followed a similar workflow as described in the "Generation of integrated SMART-seq2 based ENS atlas of the whole intestine section" above. Based on expression of known lineage markers of epithelial (*Krt8*, *Epcam*), stromal (*Pdgfra*, *Pecam1*, *Acta2*, *S100b*, *Rgs5*) and immune populations (*Cd79a*, *Mzb1*, *Cd3d*, *Trac*, *C1qa*, *Tpsab1*), the clusters were then subdivided into these three major compartments for subsequent rounds of iterative clustering and analysis. For compartment-specific clustering analysis, batch correction was performed using the following covariates: Study + Sample + Chemistry. Further quality control was done post hoc after iterative rounds of clustering to remove clusters with low UMI counts and high mitochondrial fraction. Markers used to annotate cell types in each compartment have been provided

in fig. S5, D to F. Overall this approach generated an atlas of 54 distinct non-neuronal cell types in the intestine.

Defining non-neuronal markers: The Wilcoxon rank sum test was first performed on the pan-intestine whole-tissue scRNA-seq atlas to define the markers specific to each lineage group (epithelial, stroma, endothelial, lymphocytic and myeloid). Non-neuronal, lineage-specific genes were then defined using the following criteria: lineage-specific log fold change >3, % of group expression >25%, % of reference group expression <10%. To identify genes that are shared between neuronal and non-neuronal lineages, neuron-specific markers were further derived from previously published MIRACL-seq data (*18*). This was done because we were not able to identify neurons in our pan-intestine whole-tissue scRNA-seq atlas. Finally, the neuron-enriched markers were excluded from the above list of non-neuronal, lineage-enriched genes.

Defining sex-specific markers: Differential expression analysis was performed using the MAST implementation in Seurat. Sex specific genes were defined as: Adjusted $P < 0.05$ and difference of % of expression between the two groups >10%.

PCA analysis on pseudobulk profiles: RSEM counts obtained for SMART-seq2 atlas were first rounded off. Pseudobulk analysis was performed using the decoupler python package. For region-specific analysis, pseudobulk profiles were obtained for each mouse by summing up the gene counts across all the cells obtained from the mouse. Only mice with a minimum of 10 cells in the final atlas were considered for further analysis. Filtering was performed to remove genes with a count of less than 10 across all samples. Log CPM normalization was performed and the effect of total count per cell and the number of cells contributing to the pseudobulk profile was regressed out. PCA analysis was performed on the scaled data and the first two PCs were visualized.

For assessing the effects of microbial colonization, pseudobulk profiles were obtained for each cell type in each mouse by summing up the gene counts of all cells in each condition (GF, BT, ASF or SPF). The analysis was restricted to data generated in the current study to minimize the effects of study-specific differences. Only pseudobulk profiles with a minimum of five cells were considered for further analysis. Filtering was performed to remove genes with a count of less than 10 across all samples. Log CPM normalization was performed, and the dataset was subset to the top 2000 highly variable genes. The effects of total count per cell and the number of cells contributing to the pseudobulk profile were also regressed. PCA analysis was performed on the scaled data and the first two PCs were visualized.

Differential expression analysis: Differentially expressed genes were identified using the MAST implementation in Seurat, adjusting for covariates by including them in the latent.vars parameter in the FindAllMarkers function.

To identify cluster-specific genes (CSGs) we regressed out the effects of the following covariates: Sex + Study + Condition + Mean scaled Total count + Mean scaled Epcam count. Genes with adjusted $P < 0.05$ were considered significant.

To identify region-specific genes (RSGs) we regressed out the effect of the following covariates: Cell type + Sex + Condition + Mean scaled Total count + Mean scaled Epcam count. Genes with adjusted $P < 0.05$ were considered significant.

To identify developmentally imprinted genes (born GF vs. born SPF), the following variables were adjusted: Cell type + Sex + Mean scaled Total count + Mean scaled Epcam count. Genes with adjusted $P < 0.05$ were considered significant.

To identify pan-neuronal allergy induced transcriptional response (OVAc/OVAs vs. PBS/SPF), or infection induced response (HP^+ vs. HP^-), cells from both the DJ and ileum were combined. The effects of following variables were adjusted: Sex + Cell type + Mean scaled Total count + Mean scaled Epcam count. Genes with adjusted $P < 0.1$ were considered significant. Given the lower number of cells for this comparison, we relaxed the null hypothesis criteria.

To identify cell type–specific allergy induced transcriptional response (OVAc/OVAs vs. PBS/SPF), or infection induced response (HP^+ vs. HP^-), cells from both the DJ and ileum were combined. The effects of following variables were adjusted: Sex + Cell type + Mean scaled Total count + Mean scaled Epcam count. Genes with adjusted $P < 0.2$ were considered significant. Given the lower number of cells for this comparison, we relaxed the null hypothesis criteria. The analysis was only performed for PSN Nmu-hi and the broader PEMN and PIMN cell types.

To identify pan-neuronal type 2 immune neuronal response (OVAc +OVAs + HP^+ vs. SPF + HP^-), cells from both the DJ and ileum were combined. The following covariates were adjusted: Sex + Cell type + Mean scaled Total count + Mean scaled Epcam count + Site. Genes with adjusted $P < 0.05$ were considered statistically significant.

To identify type 2 responsive genes in non-neuronal cell types, cells from the meta-analysis dataset were grouped into the following conditions: (i) OVA treated (OVAc), (ii) day 3 or day 4 after HP infection (Early HP), (iii) day 10 or day 14 after HP infection (Late HP), and (iv) All WT and basal condition samples as controls. Wilcoxon test was performed using control cells as reference for each non-neuronal cell type. For comparisons with HP infections, there were fewer cells from the immune and stromal compartment, hence the analysis was restricted to cell types from the epithelial compartment. Likewise, for comparison with OVAc there were fewer cells from the epithelial and stromal compartment, hence the analysis was restricted to cell types from the Immune compartment. Genes with adjusted $P < 0.05$ were considered significant.

For all downstream analyses based on differential comparisons, non-neuronal lineage-specific genes and sex-specific genes were removed. Gene ontology (GO) biological process and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (*69*).

Region-specific TF activity analysis: Region-specific DE genes were obtained as described above. Adjusted *P* values were first transformed to $-\log_{10}$ scale. For genes with logFC <0, the sign of transformed *P* values was set to negative and then used as inputs for estimating TF activity using the decoupler python implementation. Only TFs whose activity estimates were significant ($P < 0.05$) in any one of the regions were visualized.

Identification of microbial load–specific genes: Only PSN *Nmu*-hi cells in the ileum were considered for this analysis and further restricted to the top 1000 highly variable genes in these cells. A fixed effects generalized linear model framework was implemented to identify genes associated with the microbial load gradient. The conditions were encoded as a numeric variable (GF=1, BT=2, ASF=3 and SPF=4) to transform it from a categorical variable into a gradient series. The expression of each gene was modeled as response to the linear predictor microbial load gradient. The effects of following variables were adjusted: Mean scaled Total count + Sex + Mean scaled Epcam count + Site. Benjamini Hochberg adjusted $P < 0.1$ were considered significant. Non-neuronal lineage-specific genes and sex-specific genes were removed. The analysis was performed using the mgcv package implemented in R.

Cell proportion analysis: To identify changes in the proportions of cells across intestinal sites and between conditions, we used the miloR package (*70*). In general, this approach first constructs a KNN graph of partially overlapping representative neighborhoods instead of using a discretized cluster. A neighborhood is defined as a group of cells

connected to an index cell by an edge in the graph. Next, the number of cells in each neighborhood and each sample are counted and differential abundance analysis is performed using a negative binomial general linear model framework. Importantly for our purposes, this allows us to incorporate complex designs and adjust for covariates. Specifically, to model the proportion changes across intestinal location, we constructed a model matrix to contrast each intestinal site to the other two [for example: Colon - (Ileum + Duodenum_Jejunum)/2]. We also adjusted for the covariates Sex + Condition + Study to get robust estimates of site-enriched neighborhoods. To identify site-specific changes in cell proportions brought about by the microbiome, we compared SPF cells vs. GF-born (GF, BT and ASF) cells, adjusting for Sex using the GLM framework. Likewise, to identify proportion changes in type 2 inflammation conditions, we grouped the OVAc, OVAs and HP^+ conditions into type 2, and conversely, SPF and HP^- conditions were grouped into Controls. Since the OVAc, OVAs and SPF data were generated at a different lab site compared to the HP data, the composition analysis was performed by adjusting for Sex + Lab Site. Visualization of cell proportions was performed using the plotNhoodGraphDA function provided by the miloR package. Neighborhoods with spatial FDR $P < 0.2$ were considered significant. Additionally, density plots of neuronal subset specific UMAP embeddings were also visualized using the sc.tl.embedding_density implemented in the scanpy package (*71*). For fig. S10, F and G, cell proportions were visualized using the scCODA package (*72*).

For the non-neuronal atlas, composition analysis was restricted to GF and SPF cells. Because of the higher number of cells in the dataset, we did not estimate the spatially corrected FDR *P* values, as this would have been computationally intensive. Instead, we considered the more conservative Benjamini & Hochberg FDR $P < 0.1$ as a threshold to detect significantly different neighborhoods.

PIMN specific velocity pseudotime analysis: RNA velocity analysis was performed on a subset of 1298 PIMN sub-type neuronal cells, for which the UMAP embeddings were recomputed. Next, count matrices for spliced and unspliced reads were generated using the velocyto software's "run_smartseq2" command mapping cell specific .bam files to the *mm10* genome annotation file. The output loom files were merged and the proportion of spliced/unspliced reads for each PIMN cell subset was visualized using the scVelo package. Further filtering and normalization of the data was performed to estimate the RNA velocities using the dynamical model implementation. Finally, based on the velocity graph embedding, the latent velocity pseudotime metric was computed and visualized.

In vivo Perturb-seq library design: We considered three general criteria for selection of target genes for in vivo screening, (i) PSN *Nmu*-hi cell-specific type 2 modulators: PSN *Nmu*-hi cell type–specific changes in gene expression were considered. Genes that were either enriched in the type 2 conditions (OVAc/OVAs/HP^+) or control conditions (PBS-treated/SPF/HP^-) were selected. (ii) Pan-neuronal type 2 response regulators: Genes that were either enriched for expression in OVAc/OVAs vs. SPF comparison or HP^+ vs. HP^- comparison or both at the Pan-neuronal level. (iii) PIMN differentiation regulators: Genes whose expression changed with the differentiation trajectory from *Bglap*-hi (early progenitors) to mid-stage (*Cpne7*-hi/*Epha8*-hi) to late-stage mature neurons (*Dcc*-hi/*Aldha13*-hi). We also prioritized TFs whose expression and estimated regulon activity were associated with the differentiation trajectory. In general, we prioritized genes that were significantly enriched in the comparisons described above. However, we also selected genes that showed a trend but did not meet the significance thresholds. DE analysis and thresholds are described in the "Differential expression analysis" section.

Although we used these guidelines to design the library, we found that the prioritized genes based on a given criterion could also be selected based on another criterion. For example, expression of Tbx20 and Zfp804b were increased in the pan-neuronal type 2 response but also changed with the PIMN differentiation trajectory. In general, we utilized the library as a whole to screen for regulators of the three aforementioned biological processes.

Processing and analysis of droplet-based in vivo Perturb-seq single-cell data: FASTQ files for transcriptome library and the gRNA feature barcode library were generated using the bcl-convert software provided by Illumina. They were then aligned to the mm10 reference transcriptome using the "cellranger count" command from Cell Ranger v7.2 to generate gene expression matrices for each sample. Simultaneously, the gRNA feature barcodes were quantified by providing the protospacer and the anchor sequence for each unique gRNA as inputs to the "feature-ref" flag in the Cell Ranger pipeline. Filtered count matrices were used for further analysis using scanpy package. Cells expressing <100 detected genes were filtered out as a first-pass QC measure. In general, the downstream analysis followed a similar workflow as described in the "Generation of integrated SMART-seq2–based ENS atlas of the whole intestine" section above. Batch correction was performed using the following covariates: Experiment Batch + Sex. Further quality control was done post hoc after iterative rounds of clustering to remove clusters with low UMI counts, high mitochondrial fraction, or expression of non-neuronal lineage markers. We further removed cells from one experiment in which we observed lower UMI count and high mitochondrial fraction, leaving us with a final dataset compiled from 7 experiments. Annotation of clusters was carried out by inspecting the expression of neuronal markers described in the "Annotation and characterization of SMART-seq2–based ENS clusters" section above. Cell type–specific signatures derived from the SMART-seq2–based ENS atlas were also computed for each cluster to verify the annotations (fig. S12E).

Assigning cells with multiple gRNAs to a dominant gRNA: Multiple gRNAs (>1) were detected in 5777 cells. We reasoned that in cells with multiple gRNAs, we could identify the gRNA with the most dominant effect to increase the total number of cells in which the perturbation effect could be attributed to a single gRNA. To do so, we selected the gRNA with the highest abundance (measured as guide barcode count) in each cell and set the criterion that the abundance ratio of the gRNA relative to the sum of the rest of the gRNAs should be greater than fivefold. Using this strategy, we were able to recover a further 1178 cells in which we determined that the effect from perturbation of a single gRNA could be contributing to its transcriptional phenotype.

Cell type composition analysis on Perturb-seq data: To identify perturbations inducing a significant change, we employed a sample-size-dependent permutation approach to generate a background distribution and then computed the Mahalanobis distance metric of each perturbation to the distribution. First, we summed up the barcode counts of both replicate gRNAs per target and neuronal cell type to generate a count data matrix. Only target-level perturbations observed in >30 cells were considered for further analysis. We then transformed the count data to a proportion matrix normalized for each target perturbation and performed PCA analysis on the composition data. The first two PCs were used for visualization (Fig. 6C). Given the generally low number of total perturbations observed in our dataset, we chose to perform the analysis in the latent space to detect subtle changes in the cell type composition. Supporting our analysis approach, we find that the rotations of the non-targeting control perturbations are located approximately near the center of the distribution suggesting no major changes in composition (Fig. 6C). To account for the fact that lower cell numbers may create more noise in the cell composition and determine statistically relevant compositional changes that account for the dependence on total cell size, we implemented a permutation

approach to generate a background distribution of rotations respective to the total sample size of each perturbation. For each target gene perturbation with an observed total number of cells (suppose X), we randomly sampled the same number of cells (X) from the total dataset 10,000 times and performed PCA on the cell type composition matrix of the permuted dataset. Next, we computed the Mahalanobis distance of the first two PC rotations of the target guide of interest with reference to the permuted distribution of PC rotations. Finally, we assessed the significance of the distance using a cumulative distribution function for the chi-squared distribution using two degrees of freedom, with $P < 0.05$ considered significant.

Inference of neuronal type 2 inflammation response regulators from Perturb-seq data: Pan-neuronal type 2 response signatures were derived from the ileum-specific SMART-seq2 data as described in the "Differential expression analysis" section above. To generate the type 2 inflammation induced gene signatures, we selected genes that were significantly induced with adjusted $P < 0.05$: *Saa3*, *Tbx20*, *Zfp599*, *H3f3b*, *Rrp1*, *Reg3b*, *Ndufa8*, *Malat1*, *Wbp1*, *Pisd-ps3*, *H3f3a*, *Dnahc7b*, *2610035D17Rik*, *Pabpc1*, *Tmem63b*, *Sprr2a1*, *Nlrp5-ps*, *Soga3*, *A330093E20Rik*, *Camk1d*, *Ankar*, *Arl2*, *Ring1*, *3110057O12Rik*, *Fbxo40*, *Cmss1*, *4933409K07Rik*, *Sptlc3*, *2810008D09Rik*, *9530036O11Rik*, *Zbtb16*, *Impg1*, *Map7d1*, *Gm3893*, *Hoxb3*, *Cdk5r1*. Next, we computed module scores for the entire gene set in each cell in the Perturb-seq data, using the sc.tl.score implementation in scanpy. We then performed multivariate regression analysis modeling for the effect of each perturbation on the type 2 inflammation induced (DE-Th2-UP) module scores using the non-targeting controls as the reference. We also accounted for technical factors by including the following covariates in the linear model: Experimental Batch + total UMI count per cell + %mitochondrial fraction per cell. We treated associations with $P < 0.05$ as significant for further analysis.

Inference of PIMN differentiation regulators from Perturb-seq data: To model the effect of each perturbation of PIMN cell states along the differentiation trajectory, we first implemented a consensus NMF approach to parse the PIMN cells into gene expression programs (GEP). To do so, we subset the Perturb-Seq data to PIMN cells and select the top 1000 most variable genes for factorization. We also performed batch correction on the different experimental batches using Harmony. We ran multiple values of K ranging from 5 to 11 and selected K=6 with the highest stability solution, and density threshold=0.05 to obtain the usage matrix and the gene expression program (GEP) spectra scores, which were visualized in the reclustered PIMN UMAP embedding.

Next, we performed multivariate regression analysis modeling for the effect of each perturbation on the GEP scores, using the non-targeting controls as the reference. We also accounted for technical factors by including the following covariates in the linear model: Experimental Batch + total UMI count per cell + % mitochondrial fraction per cell. We treated associations with $P < 0.05$ as significant for further analysis.

REFERENCES AND NOTES

1. N. Y. Lai *et al.*, Gut-Innervating Nociceptor Neurons Regulate Peyer's Patch Microfold Cells and SFB Levels to Mediate Salmonella Host Defense. *Cell* **180**, 33–49.e22 (2020). doi: 10.1016/j.cell.2019.11.014; pmid: 31813624
2. W. Zhang *et al.*, Gut-innervating nociceptors regulate the intestinal microbiota to promote tissue protection. *Cell* **185**, 4170–4189.e20 (2022). doi: 10.1016/j.cell.2022.09.008; pmid: 36240781
3. J. B. Furness, The enteric nervous system and neurogastroenterology. *Nat. Rev. Gastroenterol. Hepatol.* **9**, 286–294 (2012). doi: 10.1038/nrgastro.2012.32; pmid: 22392290
4. J. B. Furness, Integrated Neural and Endocrine Control of Gastrointestinal Function. *Adv. Exp. Med. Biol.* **891**, 159–173 (2016). doi: 10.1007/978-3-319-27592-5_16; pmid: 27379644
5. S. Kulkarni *et al.*, Advances in Enteric Neurobiology: The "Brain" in the Gut in Health and Disease. *J. Neurosci.* **38**, 9346–9354 (2018). doi: 10.1523/JNEUROSCI.1663-18.2018; pmid: 30381426
6. S. Vanner, W. K. Macnaughton, Submucosal secretomotor and vasodilator reflexes. *Neurogastroenterol. Motil.* **16** (Suppl 1), 39–43 (2004). doi: 10.1111/j.1743-3150.2004.00473.x; pmid: 15066003
7. M. Rao, M. D. Gershon, Enteric nervous system development: What could possibly go wrong? *Nat. Rev. Neurosci.* **19**, 552–565 (2018). doi: 10.1038/s41583-018-0041-0; pmid: 30046054
8. Y. Obata *et al.*, Neuronal programming by microbiota regulates intestinal physiology. *Nature* **578**, 284–289 (2020). doi: 10.1038/s41586-020-1975-8; pmid: 32025031
9. P. A. Muller *et al.*, Microbiota-modulated CART$^+$ enteric neurons autonomously regulate blood glucose. *Science* **370**, 314–321 (2020). doi: 10.1126/science.abd6176; pmid: 32855216
10. A. Petsakou *et al.*, Cholinergic neurons trigger epithelial Ca^{2+} currents to heal the gut. *Nature* **623**, 122–131 (2023). doi: 10.1038/s41586-023-06627-y; pmid: 37722602
11. A. Jarret *et al.*, Enteric Nervous System-Derived IL-18 Orchestrates Mucosal Barrier Immunity. *Cell* **180**, 50–63.e12 (2020). doi: 10.1016/j.cell.2019.12.016; pmid: 31923399
12. K. M. Schneider *et al.*, The enteric nervous system relays psychological stress to intestinal inflammation. *Cell* **186**, 2823–2838.e20 (2023). doi: 10.1016/j.cell.2023.05.001; pmid: 37236193
13. J. A. Griffiths *et al.*, Peripheral neuronal activation shapes the microbiome and alters gut physiology. *Cell Rep.* **43**, 113953 (2024). doi: 10.1016/j.celrep.2024.113953; pmid: 38517896
14. D. Shi, P. Reddy, C. Walker, M. Marrin, C. Siu, N. Sharma, Differential control of intestine function by genetically defined enteric neurons. bioRxiv 2024.05.02.592229 [Preprint] (2024). https://doi.org/10.1101/2024.05.02.592229.
15. C. S. N. Klose *et al.*, The neuropeptide neuromedin U stimulates innate lymphoid cells and type 2 inflammation. *Nature* **549**, 282–286 (2017). doi: 10.1038/nature23676; pmid: 28869965
16. V. Cardoso *et al.*, Neuronal regulation of type 2 innate lymphoid cells via neuromedin U. *Nature* **549**, 277–281 (2017). doi: 10.1038/nature23469; pmid: 28869974
17. Y. Li *et al.*, Neuromedin U programs eosinophils to promote mucosal immunity of the small intestine. *Science* **381**, 1189–1196 (2023). doi: 10.1126/science.ade4177; pmid: 37708282
18. E. Drokhlyansky *et al.*, The Human and Mouse Enteric Nervous System at Single-Cell Resolution. *Cell* **182**, 1606–1622.e23 (2020). doi: 10.1016/j.cell.2020.08.003; pmid: 32888429
19. A. Dixit *et al.*, Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **167**, 1853–1866.e17 (2016). doi: 10.1016/j.cell.2016.11.038; pmid: 27984732
20. C. B. Wiese, N. Fleming, D. P. Buehler, E. M. Southard-Smith, A Uchl1-Histone2BmCherry:GFP-gpi BAC transgene for imaging neuronal progenitors. *Genesis* **51**, 852–861 (2013). doi: 10.1002/dvg.22716; pmid: 24123561
21. R. De Pace *et al.*, Messenger RNA transport on lysosomal vesicles maintains axonal mitochondrial homeostasis and prevents axonal degeneration. *Nat. Neurosci.* **27**, 1087–1102 (2024). doi: 10.1038/s41593-024-01619-1; pmid: 38600167
22. N. T. Porter, A. S. Luis, E. C. Martens, Bacteroides thetaiotaomicron. *Trends Microbiol.* **26**, 966–967 (2018). doi: 10.1016/j.tim.2018.08.005; pmid: 30193959
23. M. Wymore Brand *et al.*, The Altered Schaedler Flora: Continued Applications of a Defined Murine Microbial Community. *ILAR J.* **56**, 169–178 (2015). doi: 10.1093/ilar/ilv012; pmid: 26323627
24. J. L. Round, S. K. Mazmanian, The gut microbiota shapes intestinal immune responses during health and disease. *Nat. Rev. Immunol.* **9**, 313–323 (2009). doi: 10.1038/nri2515; pmid: 19343057
25. E. B. Brandt *et al.*, Mast cells are required for experimental oral allergen-induced diarrhea. *J. Clin. Invest.* **112**, 1666–1677 (2003). doi: 10.1172/JCI19785; pmid: 14660743
26. A. M. Mowat, W. W. Agace, Regional specialization within the intestinal immune system. *Nat. Rev. Immunol.* **14**, 667–685 (2014). doi: 10.1038/nri3738; pmid: 25234148
27. S. Picelli *et al.*, Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014). doi: 10.1038/nprot.2014.006; pmid: 24385147
28. K. Morarach *et al.*, Diversification of molecularly defined myenteric neuron classes revealed by single-cell RNA sequencing. *Nat. Neurosci.* **24**, 34–46 (2021). doi: 10.1038/s41593-020-00736-x; pmid: 33288908
29. L. Ye *et al.*, Enteroendocrine cells sense bacterial tryptophan catabolites to activate enteric and vagal neuronal pathways. *Cell Host Microbe* **29**, 179–196.e9 (2021). doi: 10.1016/j.chom.2020.11.011; pmid: 33352109
30. A. I. Sayegh, The role of bombesin and bombesin-related peptides in the short-term control of food intake. *Prog. Mol. Biol. Transl. Sci.* **114**, 343–370 (2013). doi: 10.1016/B978-0-12-386933-3.00010-8; pmid: 23317790
31. K. M. Sanders, S. M. Ward, S. D. Koh, Interstitial cells: Regulators of smooth muscle function. *Physiol. Rev.* **94**, 859–907 (2014). doi: 10.1152/physrev.00037.2013; pmid: 24987007
32. N. P. Hyland, J. F. Cryan, Microbe-host interactions: Influence of the gut microbiota on the enteric nervous system. *Dev. Biol.* **417**, 182–187 (2016). doi: 10.1016/j.ydbio.2016.06.027; pmid: 27343895
33. Z. Hu *et al.*, Small proline-rich protein 2A is a gut bactericidal protein deployed during helminth infection. *Science* **374**, eabe6723 (2021). doi: 10.1126/science.abe6723; pmid: 34735226

34. R. M. Barilla *et al.*, Type 2 cytokines act on enteric sensory neurons to regulate neuropeptide-driven host defense. *Science* **389**, 260–267 (2025). doi: 10.1126/science.adn9850; pmid: 40403128
35. V. K. Yadav, J. M. Berger, P. Singh, P. Nagarajan, G. Karsenty, Embryonic osteocalcin signaling determines lifelong adrenal steroidogenesis and homeostasis in the mouse. *J. Clin. Invest.* **132**, e153752 (2022). doi: 10.1172/JCI153752; pmid: 34905510
36. T. Laboute *et al.*, Orphan receptor GPR158 serves as a metabotropic glycine receptor: mGlyR. *Science* **379**, 1352–1358 (2023). doi: 10.1126/science.add7150; pmid: 36996198
37. L. Khrimian *et al.*, Gpr158 mediates osteocalcin's regulation of cognition. *J. Exp. Med.* **214**, 2859–2873 (2017). doi: 10.1084/jem.20171320; pmid: 28851741
38. A. Böttcher *et al.*, Non-canonical Wnt/PCP signalling regulates intestinal stem cell lineage priming towards enteroendocrine and Paneth cell fates. *Nat. Cell Biol.* **23**, 23–31 (2021). doi: 10.1038/s41556-020-00617-2; pmid: 33398177
39. S. M. Parigi *et al.*, The spatial transcriptomic landscape of the healing mouse intestine following damage. *Nat. Commun.* **13**, 828 (2022). doi: 10.1038/s41467-022-28497-0; pmid: 35149721
40. R. E. Niec *et al.*, Lymphatics act as a signaling hub to regulate intestinal stem cell activity. *Cell Stem Cell* **29**, 1067–1082.e18 (2022). doi: 10.1016/j.stem.2022.05.007; pmid: 35728595
41. B. Palikuqi *et al.*, Lymphangiocrine signals are required for proper intestinal repair after cytotoxic injury. *Cell Stem Cell* **29**, 1262–1272.e5 (2022). doi: 10.1016/j.stem.2022.07.007; pmid: 35931034
42. L. Xie *et al.*, Robust Colonic Epithelial Regeneration and Amelioration of Colitis via FZD-Specific Activation of Wnt Signaling. *Cell. Mol. Gastroenterol. Hepatol.* **14**, 435–464 (2022). doi: 10.1016/j.jcmgh.2022.05.003; pmid: 35569814
43. K. J. Jarick *et al.*, Non-redundant functions of group 2 innate lymphoid cells. *Nature* **611**, 794–800 (2022). doi: 10.1038/s41586-022-05395-5; pmid: 36323785
44. Y.-C. Wang *et al.*, Intestinal cell type-specific communication networks underlie homeostasis and response to Western diet. *J. Exp. Med.* **220**, e20221437 (2023). doi: 10.1084/jem.20221437; pmid: 36880999
45. M. A. Scavuzzo *et al.*, Enteric glial hub cells coordinate intestinal motility. bioRxiv 2023.06.07.544052 [Preprint]. (2023). https://doi.org/10.1101/2023.06.07.544052.
46. A. L. Haber *et al.*, A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017). doi: 10.1038/nature24489; pmid: 29144463
47. H. Xu *et al.*, Transcriptional Atlas of Intestinal Immune Cells Reveals that Neuropeptide α-CGRP Modulates Group 2 Innate Lymphoid Cell Responses. *Immunity* **51**, 696–708.e9 (2019). doi: 10.1016/j.immuni.2019.09.004; pmid: 31618654
48. T. Nakata *et al.*, Genetic vulnerability to Crohn's disease reveals a spatially resolved epithelial restitution program. *Sci. Transl. Med.* **15**, eadg5252 (2023). doi: 10.1126/scitranslmed.adg5252; pmid: 37878672
49. M. Biton *et al.*, T Helper Cell Cytokines Modulate Intestinal Stem Cell Renewal and Differentiation. *Cell* **175**, 1307–1320.e22 (2018). doi: 10.1016/j.cell.2018.10.008; pmid: 30392957
50. D. G. R. Tervo *et al.*, A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. *Neuron* **92**, 372–382 (2016). doi: 10.1016/j.neuron.2016.09.021; pmid: 27720486
51. A. J. Santinha *et al.*, Transcriptional linkage analysis with in vivo AAV-Perturb-seq. *Nature* **622**, 367–375 (2023). doi: 10.1038/s41586-023-06570-y; pmid: 37730998
52. K. Y. Chan *et al.*, Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nat. Neurosci.* **20**, 1172–1179 (2017). doi: 10.1038/nn.4593; pmid: 28671695
53. A. Moudgil *et al.*, Self-Reporting Transposons Enable Simultaneous Readout of Gene Expression and Transcription Factor Binding in Single Cells. *Cell* **182**, 992–1008.e21 (2020). doi: 10.1016/j.cell.2020.06.037; pmid: 32710817
54. D. Kotliar *et al.*, Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq. *eLife* **8**, e43803 (2019). doi: 10.7554/eLife.43803; pmid: 31282856
55. I. Dragoni, M. Mariotti, G. G. Consalez, M. R. Soria, J. A. Maier, EDF-1, a novel gene product down-regulated in human endothelial cell differentiation. *J. Biol. Chem.* **273**, 31119–31124 (1998). doi: 10.1074/jbc.273.47.31119; pmid: 9813014
56. K. Y. Chan *et al.*, Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nat. Neurosci.* **20**, 1172–1179 (2017). doi: 10.1038/nn.4593; pmid: 28671695
57. A. I. Asencor, G. Dvoryanchikov, V. Makhoul, P. Tsoulfas, N. Chaudhari, Selectively imaging cranial sensory ganglion neurons using AAV-PHP.S. *eNeuro* **9**, ENEURO.0373–21.2022 (2022). doi: 10.1523/ENEURO.0373-21.2022; pmid: 35610024
58. A. Wallrapp *et al.*, The neuropeptide NMU amplifies ILC2-driven allergic lung inflammation. *Nature* **549**, 351–356 (2017). doi: 10.1038/nature24029; pmid: 28902842
59. A. Laddach *et al.*, A branching model of lineage differentiation underpinning the neurogenic potential of enteric glia. *Nat. Commun.* **14**, 5904 (2023). doi: 10.1038/s41467-023-41492-3; pmid: 37737269
60. N. J. Spencer, H. Hu, Enteric nervous system: Sensory transduction, neural circuits and gastrointestinal motility. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 338–351 (2020). doi: 10.1038/s41575-020-0271-2; pmid: 32152479
61. I. S. Farooqi *et al.*, Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N. Engl. J. Med.* **348**, 1085–1095 (2003). doi: 10.1056/NEJMoa022050; pmid: 12646665
62. T. Ahrends, M. Weiner, D. Mucida, Isolation of myenteric and submucosal plexus from mouse gastrointestinal tract and subsequent flow cytometry and immunofluorescence. *STAR Protoc.* **3**, 101157 (2022). doi: 10.1016/j.xpro.2022.101157; pmid: 35146454
63. Y. Obata *et al.*, Molecular profiling of enteric nervous system cell lineages. *Nat. Protoc.* **17**, 1789–1817 (2022). doi: 10.1038/s41596-022-00697-4; pmid: 35676375
64. R. C. Challis *et al.*, Systemic AAV vectors for widespread and targeted gene delivery in rodents. *Nat. Protoc.* **14**, 379–414 (2019). doi: 10.1038/s41596-018-0097-3; pmid: 30626963
65. X. Zheng *et al.*, Massively parallel in vivo Perturb-seq reveals cell-type-specific transcriptional networks in cortical development. *Cell* **187**, 3236–3248.e21 (2024). doi: 10.1016/j.cell.2024.04.050; pmid: 38772369
66. J. M. Replogle *et al.*, Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat. Biotechnol.* **38**, 954–961 (2020). doi: 10.1038/s41587-020-0470-y; pmid: 32231336
67. I. Korsunsky *et al.*, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019). doi: 10.1038/s41592-019-0619-0; pmid: 31740819
68. L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **3**, 861 (2018). doi: 10.21105/joss.00861
69. G. Dennis Jr. *et al.*, DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **4**, P3 (2003). doi: 10.1186/gb-2003-4-5-p3; pmid: 12734009
70. E. Dann, N. C. Henderson, S. A. Teichmann, M. D. Morgan, J. C. Marioni, Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* **40**, 245–253 (2022). doi: 10.1038/s41587-021-01033-z; pmid: 34594043
71. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018). doi: 10.1186/s13059-017-1382-0; pmid: 29409532
72. M. Büttner, J. Ostner, C. L. Müller, F. J. Theis, B. Schubert, scCODA is a Bayesian model for compositional single-cell data analysis. *Nat. Commun.* **12**, 6876 (2021). doi: 10.1038/s41467-021-27150-6; pmid: 34824236
73. P. Tan, adr3545_image, version v1, Zenodo (2025); https://doi.org/10.5281/zenodo.15354644.

ACKNOWLEDGMENTS

We thank A. Regev from Broad Institute and Genentech for her scientific advice, S. Aldrich and H. Kang for editorial assistance with the manuscript and figures, C. McCabe for project management, X. Jin and X. Zheng from Scripps Research for AAV-HA-hypPB plasmid and technical advice in Perturb-seq experiments, C. Umana and A. Discua from the Broad Institute comparative medicine facility for the breeding and managing of the gnotobiotic animals, and technical support from Broad Institute flow cytometry facility and Broad Institute Genomics Platform. Schematic diagrams (https://biorender.com/6x6qq9p) and the graphical abstract (https://biorender.com/4l2ii7h) were created with BioRender.com with the indicated licenses. **Funding:** This work was supported by the Leona M. and Harry B. Helmsley Charitable Trust (to R.J.X.), the Crohn's and Colitis Foundation (to R.J.X.), a grant from the Food Allergy Research & Education (FARE) and the Food Allergy Science Initiative (to J.D. and R.J.X.), the NIH (DK135492 and DK043351, both to R.J.X.), and the Klarman Cell Observatory (to R.J.X.). **Author contributions:** Conceptualization: P.T., A.J., J.D., R.J.X.; Data curation: A.J., J.D.; Formal analysis: A.J.; Funding acquisition: J.D., R.J.X.; Investigation: P.T., S.P.M., E.M.B., H.W., C.-W.S., E.P.F., G.J.J.; Methodology: P.T., S.P.M., E.P.F.; Project administration: H.N.S., D.B.G., T.M.D., J.D., R.J.X.; Resources: H.N.S., D.B.G., T.M.D.; Software: A.J.; Validation: P.T., A.J.; Writing – original: P.T., A.J., J.D., R.J.X.; Writing – review & editing: P.T., A.J., J.D., R.J.X., with input from all other authors; Visualization: P.T., A.J.; Supervision: J.D., R.J.X. **Competing interests:** R.J.X. is co-founder of Jnana Therapeutics and Convergence Bio, Board Director at MoonLake Immunotherapeutics, consultant to Nestlé, and a member of Magnet Biomedicine and Arena Bioworks' scientific advisory boards; J.D. is a member of Biorender's scientific advisory board; these organizations had no role in this study. All other authors declare that they have no competing interests. **Data and materials availability:** Raw sequencing files are accessible on NCBI's GEO database with the following accession numbers: plate-based enteric neuron single cell atlas (GSE302208), droplet-based AAV-Perturb-seq data (GSE302051) and droplet-based intestine single cell atlas (GSE301963). Processed sequencing data are available through the Broad Institute's Single Cell Portal as SCP2971. Images analyzed and quantified as part of this study are available at Zenodo (*73*). The code used to generate the figures presented here is available from GitLab.

License information: Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adr3545
Figs. S1 to S13; Tables S1 to S9; MDAR Reproducibility Checklist

Submitted 26 June 2024; resubmitted 16 May 2025; accepted 28 July 2025
10.1126/science.adr3545

RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

Diverse somatic genomic alterations in single neurons in chronic traumatic encephalopathy

Guanlan Dong†, Chanthia C. Ma†, Shulin Mao‡, Katherine Sun-Mi Brown‡, Samuel M. Naik, Gannon A. McDonough, Samadhi P. Wijethunga, Junho Kim, Samantha L. Kirkham, Diane D. Shao, Jonathan D. Cherry, Madeline Uretsky, Elizabeth Spurlock, Ann C. McKee, August Yue Huang*, Michael B. Miller*, Eunjung Alice Lee*, Christopher A. Walsh*

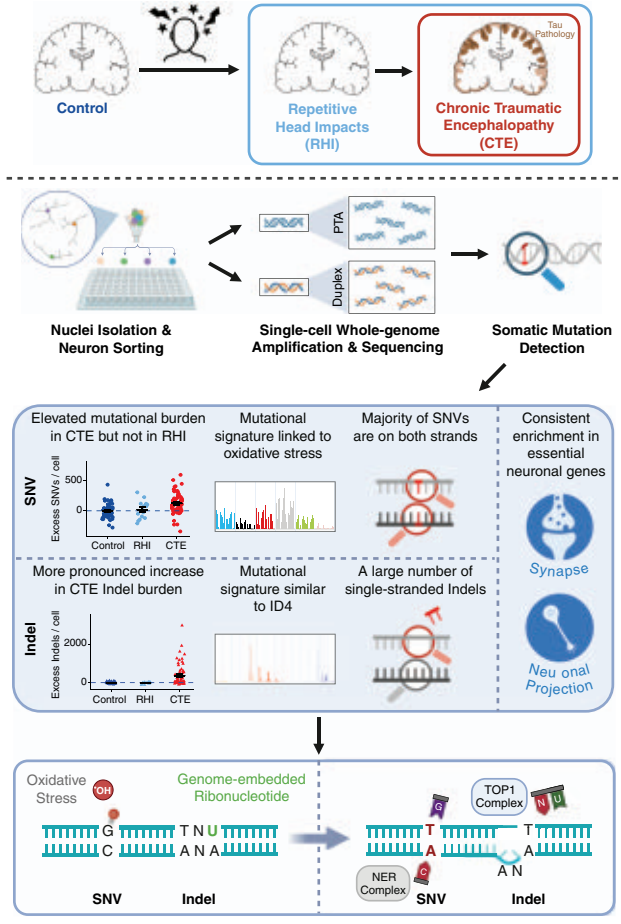
Full article and list of author affiliations: https://doi.org/10.1126/science.adu1351

INTRODUCTION: Chronic traumatic encephalopathy (CTE) is a neurodegenerative disease that develops years after exposure to repetitive head impacts (RHI) in some individuals. It is most often observed in individuals with a history of playing contact sports. Clinically, CTE is diagnosed through postmortem neuropathological examination of brain tau deposition, a hallmark it shares with Alzheimer's disease (AD). However, CTE is distinct in its tau accumulation patterns and clinical symptoms. Moreover, not everyone exposed to RHI develops CTE, and little is known about how head trauma can lead to CTE pathogenesis.

RATIONALE: While somatic mutations are well established as drivers for cancer progression, they also accumulate in noncancer cells and even in nondividing cells such as neurons. Recent advances in single-cell whole-genome sequencing (scWGS) technology have enabled the detection of private somatic mutations that accumulate in each nondividing neuron. Previous studies of single neurons showed that private somatic single-nucleotide variants (SNVs) increase during normal aging and more so in neurodegenerative diseases. However, the profiles of other mutation types have not been investigated in disease, and the role of somatic mutations in CTE is unknown.

RESULTS: Here we show that somatic mutations significantly increase in neurons of CTE patients and present specific disease-related patterns. We applied two scWGS methods to hundreds of neurons isolated from the prefrontal cortex of 15 individuals with CTE and 4 individuals with RHI but without CTE and compared mutational burdens and signatures with neurons from neurotypical controls and AD patients. In addition to the conventional strand-agnostic method, we included a strand-aware duplex sequencing modality that can distinguish between double-stranded mutations and single-stranded DNA lesions. We found a significant elevation of mostly double-stranded somatic SNVs in CTE that resembles a pattern previously reported in AD. Furthermore, we discovered high burdens of short deletions in both CTE and AD, resembling a known pattern of COSMIC signature ID4. Notably, the high burdens are only found in a subset of CTE and AD patients with a prominent signal from single-stranded events. Furthermore, somatic mutations in protein-coding genes are consistently enriched in essential neuronal functions, which, coupled with increased burdens in CTE, may lead to a dysregulated neuronal transcriptome and contribute to neurodegeneration. In contrast, somatic mutational burdens in individuals exposed to RHI but without a CTE diagnosis are similar to those of neurotypical controls, suggesting that the development of CTE involves mechanisms beyond those caused by RHI alone.

CONCLUSION: In this study, we characterized two types of somatic mutations in CTE neurons using both strand-agnostic and strand-specific methods of scWGS. Our results reveal that neurons in the CTE brain are exposed to stereotyped mutational processes shared with AD, implying potentially common pathogenic mechanisms. Specifically, increased SNVs resemble a pattern previously linked to oxidative damage, and increased short deletions resemble a pattern associated with topoisomerase 1 (TOP1) activity. In addition, the absence of similar changes in neurons exposed to RHI but without CTE highlights the complexity of CTE pathogenesis. Further study on the mechanisms leading to the short deletion pattern in certain CTE and AD patients may contribute to identifying novel therapeutic targets for neurodegenerative diseases. □



Excess of somatic mutations in single neurons in CTE. CTE develops after exposure to RHI in some individuals. Somatic mutations were profiled from single neurons in CTE brains using two scWGS methods. An excess of somatic SNVs and short deletions was found in CTE, resembling patterns associated with distinct mutagenic mechanisms that may contribute to CTE pathogenesis.

*Corresponding author. Email: christopher.walsh@childrens.harvard.edu (C.A.W.); ealice.lee@childrens.harvard.edu (E.A.L.); mbmiller@bwh.harvard.edu (M.B.M.); yue.huang@childrens.harvard.edu (A.Y.H.) †These authors contributed equally to this work. ‡These authors contributed equally to this work. **Cite this article as G. Dong *et al.*, *Science* 390, eadu1351 (2025). DOI: 10.1126/science.adu1351**

NEUROSCIENCE

Diverse somatic genomic alterations in single neurons in chronic traumatic encephalopathy

Guanlan Dong[1,2,3]†, Chanthia C. Ma[1,2,4,5]†, Shulin Mao[1,2,6]‡, Katherine Sun-Mi Brown[5]‡, Samuel M. Naik[5], Gannon A. McDonough[5], Samadhi P. Wijethunga[5], Junho Kim[7], Samantha L. Kirkham[1,2], Diane D. Shao[1,2,8], Jonathan D. Cherry[9,10,11], Madeline Uretsky[10], Elizabeth Spurlock[10], Ann C. McKee[9,10,11,12], August Yue Huang[1,2,13]*, Michael B. Miller[1,2,5,13]*, Eunjung Alice Lee[1,2,13]*, Christopher A. Walsh[1,2,13,14]*

Chronic traumatic encephalopathy (CTE) is a neurodegenerative disease linked to exposure to repetitive head impacts (RHI), yet little is known about its pathogenesis. Applying two single-cell whole-genome sequencing methods to hundreds of neurons from prefrontal cortex of 15 individuals with CTE and 4 with RHI without CTE, we revealed increased somatic single-nucleotide variants in CTE, exhibiting a pattern previously reported in Alzheimer's disease (AD). Furthermore, we discovered high burdens of somatic small insertions and deletions in a subset of CTE individuals, resembling a known pattern, ID4, also found in AD. Our results suggest that neurons in CTE experience stereotyped mutational processes shared with AD; the absence of similar changes in RHI neurons without CTE suggests that CTE involves mechanisms beyond RHI alone.

Chronic traumatic encephalopathy (CTE) develops years after exposure to repetitive head impacts (RHI) and is most often found in athletes playing contact sports (*1*–*5*). CTE is diagnosed by postmortem neuropathological examination, on the basis of the presence of a pathognomonic lesion consisting of perivascular hyperphosphorylated tau neurofibrillary tangles in neurons at the depths of the cortical sulci (*1*, *2*). Tau deposition is common to both CTE and Alzheimer's disease (AD) (*4*, *6*); however, CTE is characterized by distinct neuropathological features, tau molecular structural conformation (*7*, *8*), and clinical symptoms (*5*, *9*). The precise mechanisms by which RHI induce tau neurodegeneration are poorly understood. Somatic mutations are well-known drivers of cellular proliferation in neoplasia (*10*–*12*), but they also accumulate in nonneoplastic cells and even nondividing neurons with age (*13*–*17*). Multiple studies using single-cell whole-genome sequencing (scWGS) (*17*, *18*) and single-molecule duplex sequencing (*16*, *19*, *20*) converge on accumulation rates of 16 to 17 somatic single-nucleotide variants (sSNVs) per year and two to three somatic small insertions and deletions (sIndels) per year in neurons from neurotypical controls. Furthermore, neurons from individuals with AD and other neurodegenerative disorders show higher sSNV burden with distinct mutational signatures (*18*, *21*). In this study, we applied scWGS to amplified DNA from single neuronal nuclei isolated from CTE, RHI without CTE, AD, and neurotypical control individuals (Fig. 1A), using both strand-agnostic and strand-specific amplification methods to assess somatic mutations in CTE.

sSNVs in CTE neurons

We acquired frozen brain tissue from dorsolateral prefrontal cortex (PFC) from individuals with a history of RHI, with and without CTE, as well as neurotypical controls. We isolated nuclei of single neurons by staining for the nuclear dye 4′,6-diamidino-2-phenylindole (DAPI) and the pan-neuronal marker NeuN and sorted neurons using fluorescence-activated nuclei sorting (FANS), gating specifically for the largest NeuN-positive nuclei (Fig. 1B and fig. S1). This gating method isolates pyramidal excitatory neurons with a purity of >99% (*18*). We sorted single nuclei into individual wells of 96-well plates then performed whole-genome amplification using primary template amplification (PTA) (*22*) or multiplexed end-tagging amplification of complementary strands (META-CS) (*19*). Amplified DNA underwent multiple screening and quality control steps, ensuring that only well-amplified genomes were used for sequencing. We sequenced PTA-amplified genomes of 68 neurons from 15 cases of CTE (exposed to RHI with neuropathologically verified CTE) and 16 neurons from 4 cases of RHI (exposed to RHI without CTE) and compared these with 56 neurons from 19 neurotypical controls (*23*) and 27 neurons from 7 cases of AD (*18*) (tables S1 to S3).

We identified both sSNVs and sIndels in each single PTA neuron using SCAN2 (*17*) and found significantly increased sSNVs in CTE, averaging 114 more sSNVs per cell compared with controls [$P = 0.003$, linear mixed-effects (LME) model; Fig. 1C]. The increased sSNV burdens are comparable to those in AD (Fig. 1D). In contrast, sSNV numbers in RHI cases were indistinguishable from controls ($P = 0.725$, LME model; Fig. 1C), suggesting that increased sSNVs in CTE arise not from RHI exposure alone but from additional factors associated with the development of CTE. All CTE and RHI brain samples were collected and processed at the UNITE Brain Bank to control for potential batch effects, although the small sample size of RHI may limit statistical power. We performed nonparametric tests on somatic burdens after adjusting for age, where both CTE and AD neurons showed a significant excess of sSNVs compared with controls (CTE, $P = 3.3 \times 10^{-6}$; AD, $P = 2.3 \times 10^{-5}$; two-tailed Wilcoxon test; Fig. 1E and fig. S2, A to C). The increased sSNVs in CTE remained significant after controlling for additional quality metrics (see methods and fig. S3). Furthermore, although neurons in each case showed some variation, we observed greater mean sSNVs in nearly all CTE cases than the burden attributable to normal aging (Fig. 1F).

Somatic mutations were broadly distributed across each neuron's genome (Fig. 1G). Notably, we found no association between sSNV burden and potential risk factors for CTE pathology, including cortex topology (*1*) (fig. S4) and years of playing American football (*6*, *24*–*26*), which implies that the duration of RHI exposure does not directly lead to increased sSNVs, in line with the similar sSNV burden between RHI and controls.

To investigate the strand specificity of excess sSNVs in CTE, we used a modified duplex META-CS method (*19*) (see methods) to profile 115 neurons from 12 CTE cases and 146 neurons from 16 neurotypical controls with paired PTA data (table S1). We distinguished variants representing double-stranded mutations and single-stranded DNA lesions

[1]Division of Genetics and Genomics, Manton Center for Orphan Disease Research, Boston Children's Hospital, Boston, MA, USA. [2]Department of Pediatrics, Harvard Medical School, Boston, MA, USA. [3]Bioinformatics and Integrative Genomics Program, Harvard Medical School, Boston, MA, USA. [4]Harvard-MIT MD-PhD Program, Harvard Medical School, Boston, MA, USA. [5]Division of Neuropathology, Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [6]Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, MA, USA. [7]Department of Biological Sciences, Sungkyunkwan University, Suwon, South Korea. [8]Department of Neurology, Boston Children's Hospital, Boston, MA, USA. [9]Veterans Affairs (VA) Boston Healthcare System, US Department of Veteran Affairs, Boston, MA, USA. [10]Alzheimer's Disease Research Center and Chronic Traumatic Encephalopathy Center, Chobanian and Avedisian School of Medicine, Boston University, Boston, MA, USA. [11]Department of Pathology and Laboratory Medicine, Chobanian and Avedisian School of Medicine, Boston University, Boston, MA, USA. [12]Department of Neurology, Chobanian and Avedisian School of Medicine, Boston University, Boston, MA, USA. [13]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [14]Boston Children's Hospital, Howard Hughes Medical Institute, Boston, MA, USA. ***Corresponding author. Email: christopher.walsh@childrens.harvard.edu (C.A.W.); ealice.lee@childrens.harvard.edu (E.A.L.); mbmiller@bwh.harvard.edu (M.B.M.); yue.huang@childrens.harvard.edu (A.Y.H.)** †These authors contributed equally to this work. ‡These authors contributed equally to this work.

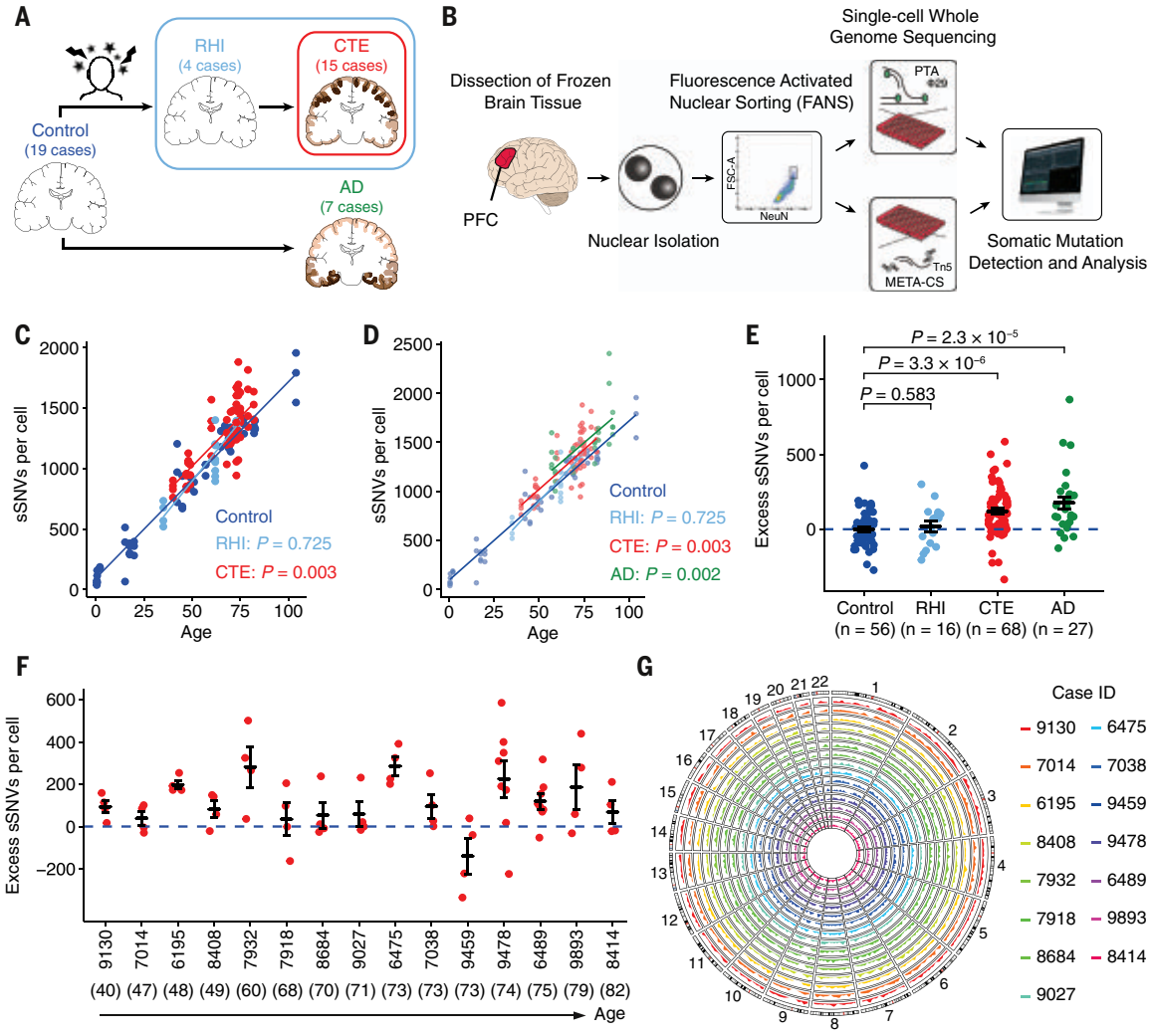


Fig. 1. Study design and sSNV burden in CTE neurons. (**A**) Cohort design: neurotypical control, repetitive head impacts (RHI), chronic traumatic encephalopathy (CTE), and Alzheimer's disease (AD). Illustrations of CTE and AD show characteristic tau pathology patterns (in brown). The number of cases included in this study is shown for each clinical condition. (**B**) Single-cell whole-genome sequencing (scWGS) experimental approach. Nuclei were isolated from postmortem human brain tissue and subjected to fluorescence-activated nuclear sorting for the neuronal nuclear marker NeuN. Nuclei were sorted, lysed, and subjected to primary template amplification (PTA) and multiplexed end-tagging amplification of complementary strands (META-CS). Amplified genomic DNA was then assayed using WGS to identify somatic mutations. PTA data were used to determine somatic mutation burden, while META-CS data were used to identify strand-related signatures. (**C**) sSNV burden in CTE (red), RHI (light blue), and neurotypical control (dark blue) brains with a significant increase in CTE but not in RHI when compared with controls. sSNV burden (from each neuron as a point) estimated by SCAN2 is fitted against age by clinical conditions using LME models (neurotypical control, dark blue; RHI, light blue, $P = 0.725$; CTE, red, $P = 0.003$). *P* values compare each clinical condition against controls. (**D**) Similar to (C), with added neurons from AD brains (green, $P = 0.002$ using the LME model). *P* values compare each clinical condition against controls. (**E**) Excess sSNV burden in RHI, CTE, and AD compared with neurotypical control after adjusting for age. Data are mean ± standard error. The dashed blue line shows sSNVs attributable to age (zero excess). *P* values are from two-tailed Wilcoxon tests. (**F**) Excess sSNVs in each CTE case ordered by increasing age (indicated in parentheses). The dashed blue line shows sSNVs attributable to age (zero excess). (**G**) Circos plot showing the PTA sSNV density distribution of CTE cases across the whole genome. Each CTE case is depicted by color in a circular track.

(fig. S5A). According to META-CS–derived strand-specific signatures, PTA-profiled sSNVs showed, on average, 85% double-stranded signature contribution (fig. S5B), indicating that PTA-identified sSNVs are predominantly double-stranded, albeit not exclusively. This aligns with the highly similar rate of age-related sSNV increase when profiled by either PTA (*17*) or strand-specific methods (*16*, *19*).

Mutational signatures of sSNVs in CTE neurons

To investigate potential sources of sSNVs in CTE neurons, we performed mutational signature analysis on the PTA scWGS data from CTE and neurotypical control neurons. First, we decomposed the sSNVs into signatures A and C, previously identified in single neurons (*21*) (Fig. 2, A and B). Signature A closely resembles SBS5 from the COSMIC database (v.3.2) (*27*), a clocklike signature associated with aging (SBS, single base substitution). Signature C resembles SBS8, a signature associated with deficient transcription-coupled nucleotide excision repair (TC-NER) (*21*, *28*), along with other SBS signatures, which have also been linked with oxidative damage in AD neurons (*18*). The contribution of signature A increased with age without showing differences between CTE and control neurons (Fig. 2, C and D). In contrast, we observed a significantly larger contribution of signature C in CTE neurons, with an average excess of 64 sSNVs per cell ($P = 7.7 \times 10^{-4}$, LME

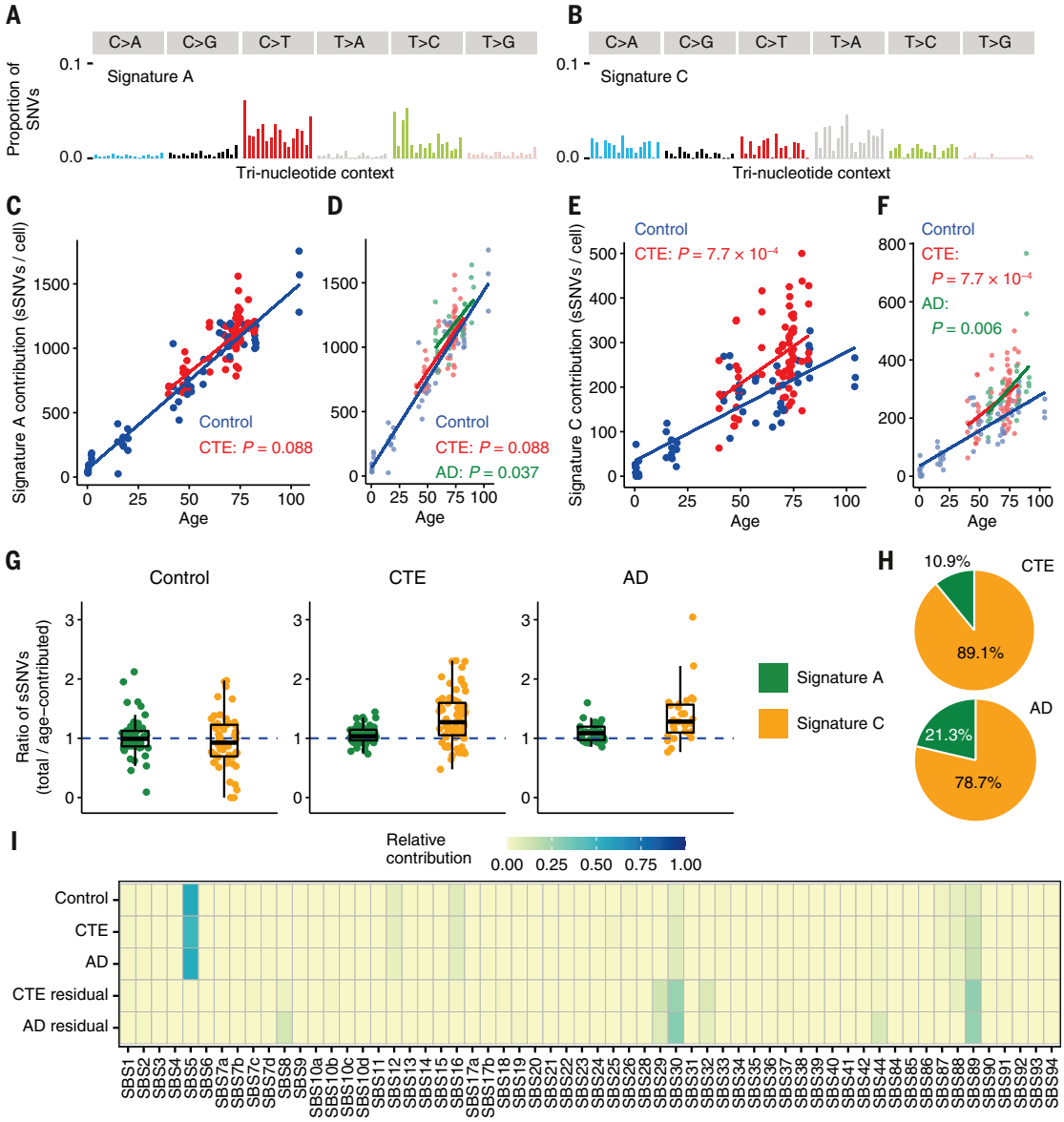


Fig. 2. sSNV mutational signatures in CTE neurons. (**A** and **B**) Mutation spectra of signatures A and C. (**C** to **F**) sSNV burden separated by contributions from signature A [(C) and (D)] and signature C [(E) and (F)] in CTE (red), AD (green), and neurotypical control (dark blue) neurons. These signatures decompose sSNV burden into age- and disease-specific effects. Their contributions are fitted against age by clinical conditions using LME models [(C), signature A, CTE: $P = 0.088$; (D), signature A, CTE: $P = 0.088$, AD: $P = 0.037$; (E), signature C, CTE: $P = 7.7 \times 10^{-4}$; (F), signature C, CTE: $P = 7.7 \times 10^{-4}$, AD: $P = 0.006$]. *P* values compare each clinical condition against controls. (**G**) Ratios of each signature's total contribution to age-related contribution in neurotypical control, CTE, and AD neurons. Age-contributed sSNVs of each signature are obtained from the LME model in neurotypical controls. A ratio of >1 indicates a higher contribution from the signature compared with age-matched controls. Bars in each box plot from top to bottom show the first, second (median), and third quartile; whiskers extend 1.5 interquartile range (IQR). (**H**) Relative contribution of signatures A and C in CTE and AD after adjusting for age. Based on the ratios shown in (G), the relative contribution of each signature is calculated by removing the median ratio of neurotypical controls (representing age effect) from the median ratio of each disease. (**I**) Mutational spectra of neurotypical control, CTE, and AD neurons are fitted to the COSMIC SBS database of cancer mutational signatures. Residual mutational patterns for CTE and AD are obtained by subtracting mutation profiles of age-matched controls from those of CTE and AD to show disease-specific contributions.

model; Fig. 2, E and F). We also compared the signature contributions to 27 previously reported AD neurons (*18*), in which AD showed a greater contribution of both signature A and signature C compared with age-matched controls ($P = 0.037$ and 0.006, respectively, LME model; Fig. 2, D and F).

To better understand the age- and disease-associated contributions of the two signatures, we calculated the ratio of each signature's total sSNVs to age-contributed sSNVs and found an elevated contribution of signature C (ratio > 1) in CTE and AD (Fig. 2G), suggesting shared mutational mechanisms in both neurodegenerative tauopathies. When we compared the relative contribution of the two signatures after adjusting for age, we found that signature C accounted for a greater proportion of the increase in CTE than in AD (Fig. 2H). META-CS data identified a mean contribution of signature C ranging from 22 to 24% in double-stranded SNVs (dsSNVs) and around 67% in single-stranded SNVs (ssSNVs) without differences between CTE and age-matched controls (fig. S6). This finding, combined with the predominant contribution of dsSNVs (~85%) in PTA neurons, suggests that a substantial amount

of signature C is double-stranded. Although the precise mechanism of signature C and SBS8-like variants has not been definitively determined, the C>A mutations in this pattern are associated with alterations in reactive oxygen species (ROS) in other contexts (*29*) and represent one possible contributor to disease pathogenesis.

We also conducted de novo signature analysis using nonnegative matrix factorization (NMF) (*30*) (see methods). We identified two de novo signatures, N1 and N2, that shared many similarities with signatures C and A, respectively, but also had substantial overlaps with each other (such as SBS5 and SBS16) (fig. S7) which may limit their interpretability as they may not represent distinct biological processes. To investigate other potential contributing factors to the excess sSNVs in CTE, we aggregated mutations on the basis of their disease status (neurotypical control, CTE, and AD) and fitted them to COSMIC SBS signatures (Fig. 2I). The aging effect represented by SBS5 dominated across all three groups. However, after subtracting the pattern of age-matched controls from CTE and AD, we obtained residual patterns that were disease specific (Fig. 2I and fig. S7D), suggesting elevated contributions of SBS29 and SBS32 in CTE that are almost absent in controls. SBS29 is a signature associated with tobacco chewing, yet distinct from tobacco smoking–related SBS4 and SBS92 (*27*). A high prevalence of smokeless tobacco usage has been reported in athletes (*31*, *32*); smokeless tobacco usage in American football players is unknown. SBS32 is associated with azathioprine treatment and was recently reported to be strongly associated with age in oligodendrocytes, with an absence in aging neurons (*23*). Other COSMIC signatures, including SBS30 and SBS89, were present in controls but increased in CTE; SBS30 contributed to signature C (fig. S7C) and was associated with deficient base excision repair which may also play a role in the excessive DNA damage, whereas the mechanism underlying SBS89 is unclear.

sIndels in CTE neurons

Compared with sSNVs, sIndels showed a more pronounced increase in CTE than in controls ($P = 0.004$, LME model; $P = 2.8 \times 10^{-8}$, two-tailed Wilcoxon test; Fig. 3, A and B, and fig. S2, D to F). RHI neurons showed no differences in sIndel burden versus controls, further supporting the hypothesis that increased somatic burden is specific to CTE pathology. CTE neurons accumulated a mean of 312 excess sIndels per cell compared with controls, approximately three times the CTE-specific sSNV increase and equivalent to sIndels accumulated over >100 years (given the normal neuronal rate of two to three sIndels per year per genome). These CTE excess sIndels are predominantly 2- to 4-bp (base pair) deletions (Fig. 3C). Reanalysis of PTA AD neurons from our previous study (*18*) revealed a similar phenomenon in AD (fig. S8).

We developed a new computational pipeline (see methods) specifically for identifying sIndels in single-cell META-CS data and extracted mutational signatures that represented double-stranded and single-stranded Indels (dsIndels and ssIndels) in CTE and neurotypical controls (Fig. 3D). We confirmed the robustness of these signatures with different calling thresholds (fig. S9, A to D). After fitting the dsIndel and ssIndel signatures to PTA-profiled sIndels (see methods), we found that the excess sIndels in some CTE individuals are primarily single-stranded, whereas other CTE individuals and controls have more modest numbers of sIndels that are primarily double-stranded (Fig. 3E).

To explore potential biological processes giving rise to dsIndels and ssIndels in CTE, we decomposed META-CS–derived signatures into COSMIC Indel (ID) signatures (v.3.2) (Fig. 3F). Some signatures were present in both CTE and neurotypical controls. ID5, a clocklike signature associated with aging, contributes to dsIndels and at a lesser degree to ssIndels of both groups. ID12, a signature with unknown etiology where Indels occur primarily in repetitive regions, showed a strong contribution to ssIndels of both groups, implying that repetitive regions may be prone to single-stranded lesions or sequencing errors. As the current technology may not distinguish sequencing errors or amplification artifacts from single-stranded lesions, we focused on differences observed between control and CTE that are more likely to reflect biological processes. ID4, mainly characterized by 2- to 4-bp deletions, has a predominant and robust (fig. S9, E to H) presence in CTE and is minimal in controls, consistent with the bias toward 2- to 4-bp deletions observed in CTE PTA neurons (Fig. 3C). Although the signal is more pronounced in the ssIndel signature, the dsIndel signature also showed a substantial ID4 contribution, indicating a mutagenic process where a subset of ID4-associated DNA lesions become double-stranded mutations.

Somatic ID4-like deletions in certain CTE individuals

The excess sIndels in CTE were driven by a subset of individuals (Fig. 4A), designated "High-Indel CTE," whereas other CTE individuals exhibited much lower excess sIndels compared with age-matched controls, designated "Low-Indel CTE." This distinction in excess sIndels between High-Indel and Low-Indel CTE groups remained after controlling for quality metrics (see methods and fig. S10). Although individuals in the High-Indel group were generally older (>70 years old), two younger individuals suggest that other factors beyond age may contribute. We observed modest correlation between excess sIndels and excess sSNVs in High-Indel CTE at the cellular level, but limited power precludes a clear conclusion (fig. S11). Furthermore, we found a significant association between sIndels in CTE and the duration of symptoms (the time between symptom onset and death; $P = 0.022$, LME model), which was stronger than the association with age at death ($P = 0.059$, LME model), suggesting that the duration of disease might contribute to excess sIndels. The apolipoprotein E (*APOE*) ε4 allele has been identified as a prominent risk factor for AD (*33*) and is associated with the severity of CTE tau pathology (*34*), however, we found no association of *APOE* ε4 with sIndels in CTE after controlling for age (see methods). Years of playing football—a proxy for cumulative RHI exposure—also did not show association, consistent with our findings in sSNVs. Reanalysis of PTA AD neurons (*18*) revealed similarly high rates of sIndels in certain AD individuals (designated "High-Indel AD" and "Low-Indel AD"; Fig. 4B), indicating that excessive sIndels may contribute to multiple tau-based neurodegenerative diseases.

The High-Indel CTE and AD individuals present a distinct mutational pattern. Using the counts of each context of PTA sIndels detected in each individual, unsupervised clustering identified High-Indel CTE and AD as a separate cluster from Low-Indel CTE and AD and neurotypical controls (Fig. 4C and fig. S12). The same clustering results were reinforced by META-CS data (fig. S13). The distinct pattern in High-Indel CTE and AD individuals is mainly characterized by 2- to 4-bp deletions, where the proportion of 2- to 4-bp deletions is significantly higher in High-Indel CTE than in age-matched controls for both dsIndels and ssIndels ($P = 1.0 \times 10^{-4}$ and $P = 7.5 \times 10^{-4}$, respectively, two-tailed Wilcoxon test, Fig. 4, D and E). In line with the fact that ID4 comprises primarily 2- to 4-bp deletions, we observed a consistent contribution of both double- and single-stranded ID4 across all High-Indel CTE (Fig. 4F), significantly higher than in Low-Indel CTE (fig. S14; dsIndel: $P = 0.003$, ssIndel: $P = 0.005$, two-tailed Wilcoxon test) and control groups (fig. S14; dsIndel: $P = 0.001$, ssIndel: $P = 1.0 \times 10^{-4}$, two-tailed Wilcoxon test). This highlights the role of ID4-associated biological processes in elevated DNA damage in CTE. Although the etiology of ID4 is not completely understood, recent studies (*35*, *36*) demonstrated that ID4 may be mediated by topoisomerase 1 (TOP1) activity and linked to neurodegeneration.

Influence of transcriptional activity and chromatin accessibility on somatic mutation in CTE

Gene transcription and epigenetic state are closely linked to DNA damage and repair in neurons (*37*, *38*). In contrast to mitotic cells,

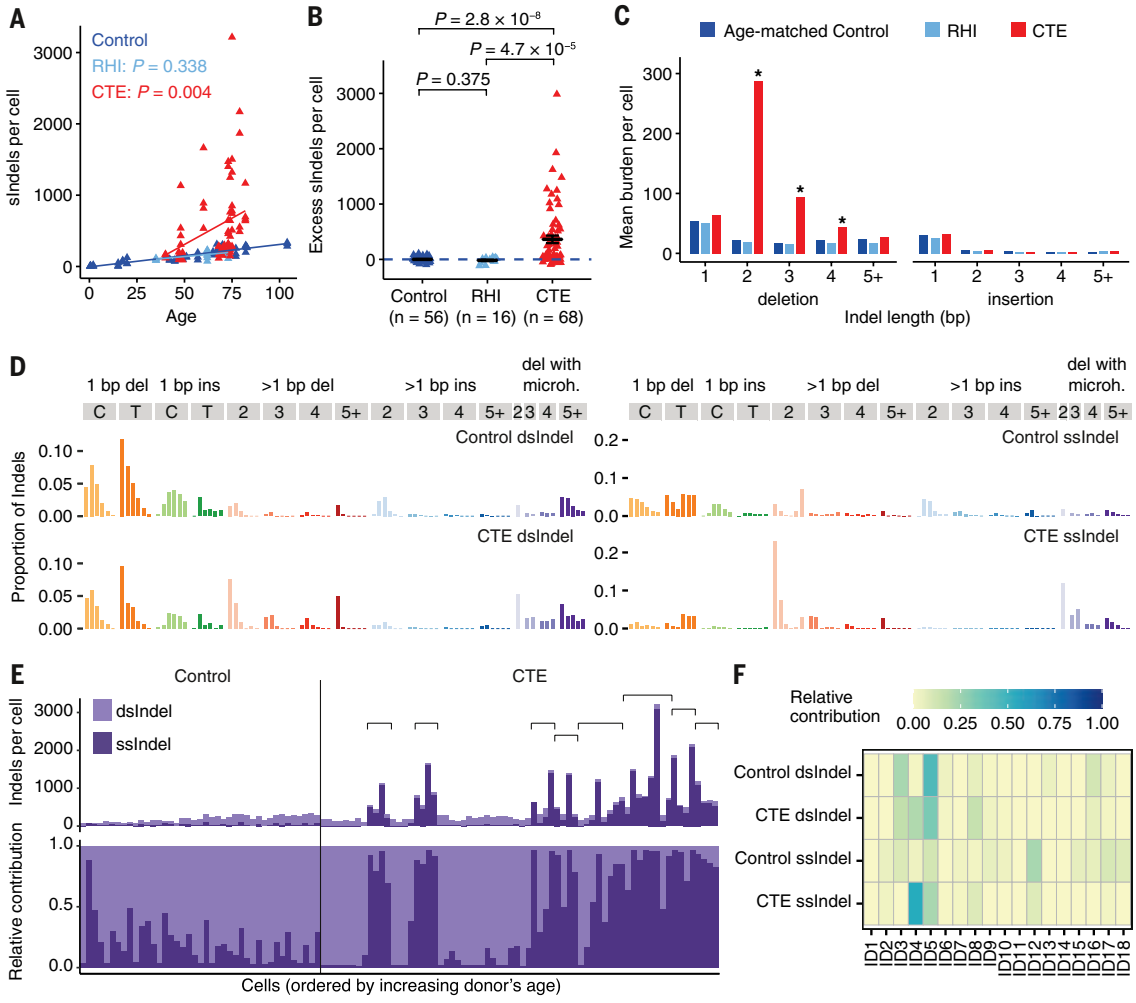


Fig. 3. sIndel burden and mutational signatures in CTE neurons. sIndel burdens were identified from PTA scWGS data, with strand-related analysis inferred from mutational signatures extracted from META-CS scWGS data. (**A**) sIndel burden in CTE (red), RHI (light blue), and neurotypical control (dark blue) brains with a significant increase in CTE but not in RHI when compared with controls. sIndel burden (from each neuron as a triangle) estimated by SCAN2 is fitted against age by clinical conditions using LME models (neurotypical control, dark blue; RHI, light blue, $P = 0.338$; CTE, red, $P = 0.004$). *P* values compare each clinical condition against controls. (**B**) Excess sIndel burden in RHI and CTE compared with neurotypical control after adjusting for age. Data are mean ± standard error. The dashed blue line shows sIndels attributable to age (zero excess). *P* values are from two-tailed Wilcoxon tests. (**C**) Comparison of all types of sIndels across age-matched controls, RHI, and CTE. Data are mean burden per cell. Asterisk denotes significant changes in certain types of sIndels when compared with age-matched controls ($P < 0.05$, two-tailed Wilcoxon test). (**D**) Double-stranded (ds) and single-stranded (ss) Indel signatures in neurotypical controls and CTE extracted from META-CS data. (**E**) Absolute (top) and relative (bottom) contribution of dsIndel and ssIndel signatures in PTA-profiled neurotypical control and CTE neurons. The pair of dsIndel and ssIndel signatures used for decomposition is determined by the clinical condition of PTA neurons. CTE cases with a pronounced ssIndel contribution are indicated by brackets. Cells with <15 Indels are not shown. (**F**) Decomposition of dsIndel and ssIndel signatures to COSMIC ID database.

replication-independent processes account for most somatic mutations in healthy neurons, where both gene expression and chromatin accessibility have been reported to be positively correlated with mutation burden (*17*, *18*, *23*). To investigate whether somatic mutations have distinct enrichment patterns in CTE neurons, we generated single-nucleus RNA sequencing (snRNA-seq) data from CTE and neurotypical control samples (fig. S15) and used published single-nucleus assay for transposase-accessible chromatin with sequencing (snATAC-seq) data from sample-matched neurotypical controls (*23*). We extracted neuronal profiles to eliminate potential biases from other cell types.

We found somatic mutations significantly more enriched in transcribed genes than in nontranscribed genes in control and CTE ($P < 0.05$ for all, two-tailed Wilcoxon test), which aligns with previous findings and highlights the role of transcriptional activity in neuronal mutagenesis. We further confirmed positive correlations of somatic mutations with gene expression and chromatin accessibility in both control and CTE (Fig. 5A). Using signatures to further dissect enrichment patterns according to mutational processes, we observed positive correlations between gene expression and sSNVs attributed to signature A for both control and CTE, whereas signature C showed a negative correlation (Fig. 5B), consistent with previous results in AD neurons (*18*) and reinforcing the shared findings between CTE and AD. Chromatin accessibility showed no clear trend for signature C–related sSNVs, which may be explained by complex relationships between chromatin accessibility and gene expression activity in highly expressed genes (fig. S16) as well as potential confounds of global genome NER (GG-NER)–related sSNVs in regulatory elements, as signature C was identified in genetic disorders with deficiency in either TC-NER or GG-NER (*21*). For sIndels, we separately extracted double- and single-stranded signatures from High-Indel CTE, Low-Indel CTE, and controls and found that total burden as well as burden attributed to either signature showed moderate positive correlations with both

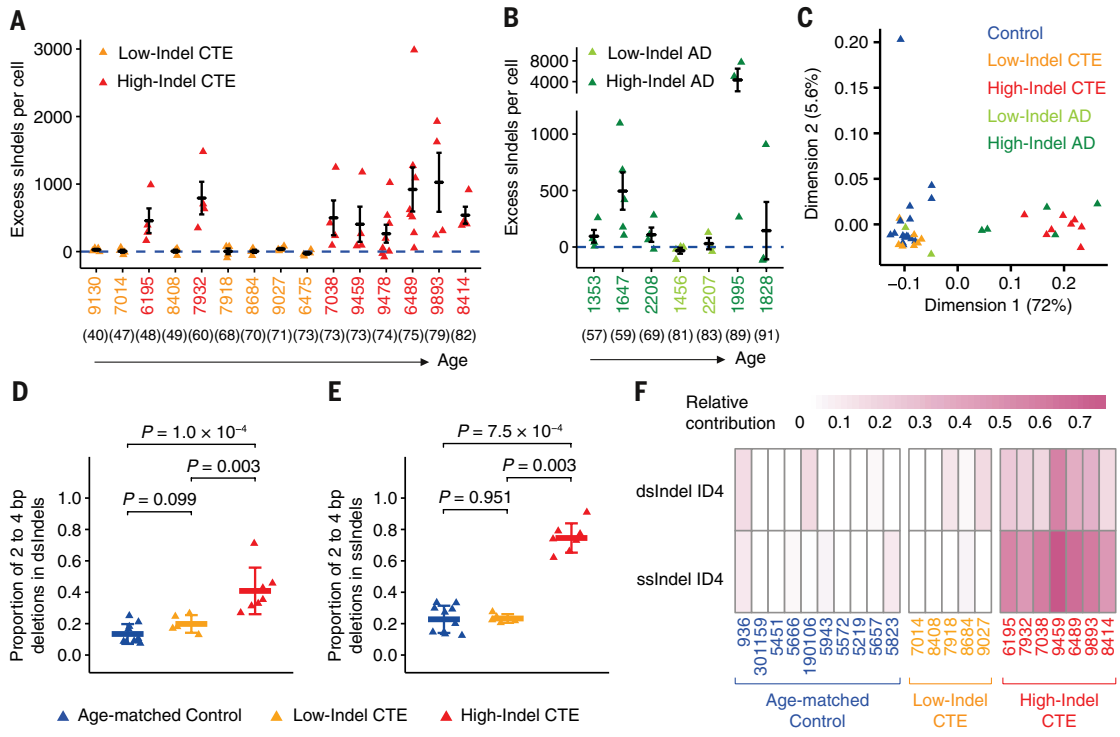


Fig. 4. Somatic ID4-like deletions in certain individuals with CTE. (**A** and **B**) Excess sIndels in each CTE (A) and AD (B) case ordered by increasing age (indicated in parentheses). CTE and AD cases are separated into either the High-Indel group (High-Indel CTE, red; High-Indel AD, dark green) or the Low-Indel group (Low-Indel CTE, yellow; Low-Indel AD, light green) on the basis of whether they have an excess of sIndels (>50). Data are mean ± standard error. The dashed blue line shows sIndels attributable to age (zero excess). (**C**) Principal components analysis (PCA) clustering of sIndels from each case across neurotypical controls, CTE, and AD. sIndels from each case were aggregated and stratified into 83 contexts defined by COSMIC. The first two dimensions are used for visualization with the percentage of variance explained shown for each dimension. Cases with <15 sIndels are not shown. (**D** and **E**) Comparison of proportions of 2- to 4-bp deletions in dsIndels (D) and ssIndels (E) across age-matched controls, Low-Indel CTE, and High-Indel CTE. Each triangle represents an individual from META-CS data. High-Indel CTE versus age-matched controls: $P = 1.0 \times 10^{-4}$ and $P = 7.5 \times 10^{-4}$, two-tailed Wilcoxon test. Data are mean ± standard deviation. (**F**) Relative contribution of ID4 to dsIndels and ssIndels of each case from META-CS data shown as a heatmap. Case IDs are colored by their group assignment (age-matched control, dark blue; Low-Indel CTE, yellow; High-Indel CTE, red).

gene expression and chromatin accessibility (Fig. 5, C and D). The similar pattern between dsIndel and ssIndel suggests that they share similar mutagenic mechanisms, with single-stranded lesions representing an intermediate stage before double-stranded mutations.

We further examined the distribution of somatic mutations in protein-coding genes for their potential functional impact. Gene Ontology (GO) analysis of genes mutated by sSNVs (Fig. 5E) and sIndels (Fig. 5F) revealed consistent enrichments in neuronal functions such as synaptic structures and neuron projection. This observation aligns with our finding that somatic mutations are enriched in regions of high expression and open chromatin, reflecting the active transcription of neuronal genes in neurons. Taken together with the CTE-specific somatic burden increase, our results suggest that excess somatic mutations, notably sIndels in CTE neurons, may frequently disrupt essential neuronal genes and eventually contribute to neurodegeneration (fig. S17).

Discussion

In this study, we characterized somatic mutations in CTE neurons using two different scWGS methods and revealed elevated DNA damage leading to sSNVs and sIndels in CTE, in patterns distinct from normal aging. We propose specific mutagenic pathways for the accumulation of sSNVs and sIndels (fig. S17). Excess sSNVs in CTE were largely contributed by signature C, a pattern that is rare in controls and elevated in AD. Given the shared feature of tau pathology, CTE and AD may share common pathways for accumulating DNA damage, possibly related to oxidative stress (*18*). We also found a subset of High-Indel CTE individuals harboring excess double- and single-stranded Indels, with a predominant contribution from ID4 (*27*). Analysis of AD individuals showed a subset with similar excessive sIndels, which was also seen in a parallel study examining tau pathologic states (*39*).

Previous studies have highlighted RHI as the likely cause of CTE (*40*, *41*); however, not all individuals exposed to RHI develop CTE. We found that the somatic mutation burden of the RHI group was similar to that of neurotypical controls, implying that the elevated DNA damage in CTE is distinct from long-term exposure to RHI. A limitation of this study is that the RHI group is small with a narrow age range, and a larger group of varying ages and RHI exposures is needed to validate this finding. Our cohort focused on severe CTE (stages III and IV); future studies that include early-stage CTE could assess the role of somatic mutations in disease progression. Moreover, other distinctive features of CTE, such as the perivascular localization of tau deposition at the depths of the cortical sulci, may reveal previously unknown associations between DNA damage and pathogenesis. Our analysis between the sulcus and gyrus showed no difference (fig. S4), but statistical power was limited. In addition, we did not detect any somatic mutations in genes known to be associated with neurodegenerative tauopathies (*APP*, *PSEN1*, *PSEN2*, *APOE*, and *MAPT*), similar to a previous report in AD (*18*).

Duplex sequencing data distinguished single-stranded and double-stranded genomic alterations, showing that sSNVs and sIndels in

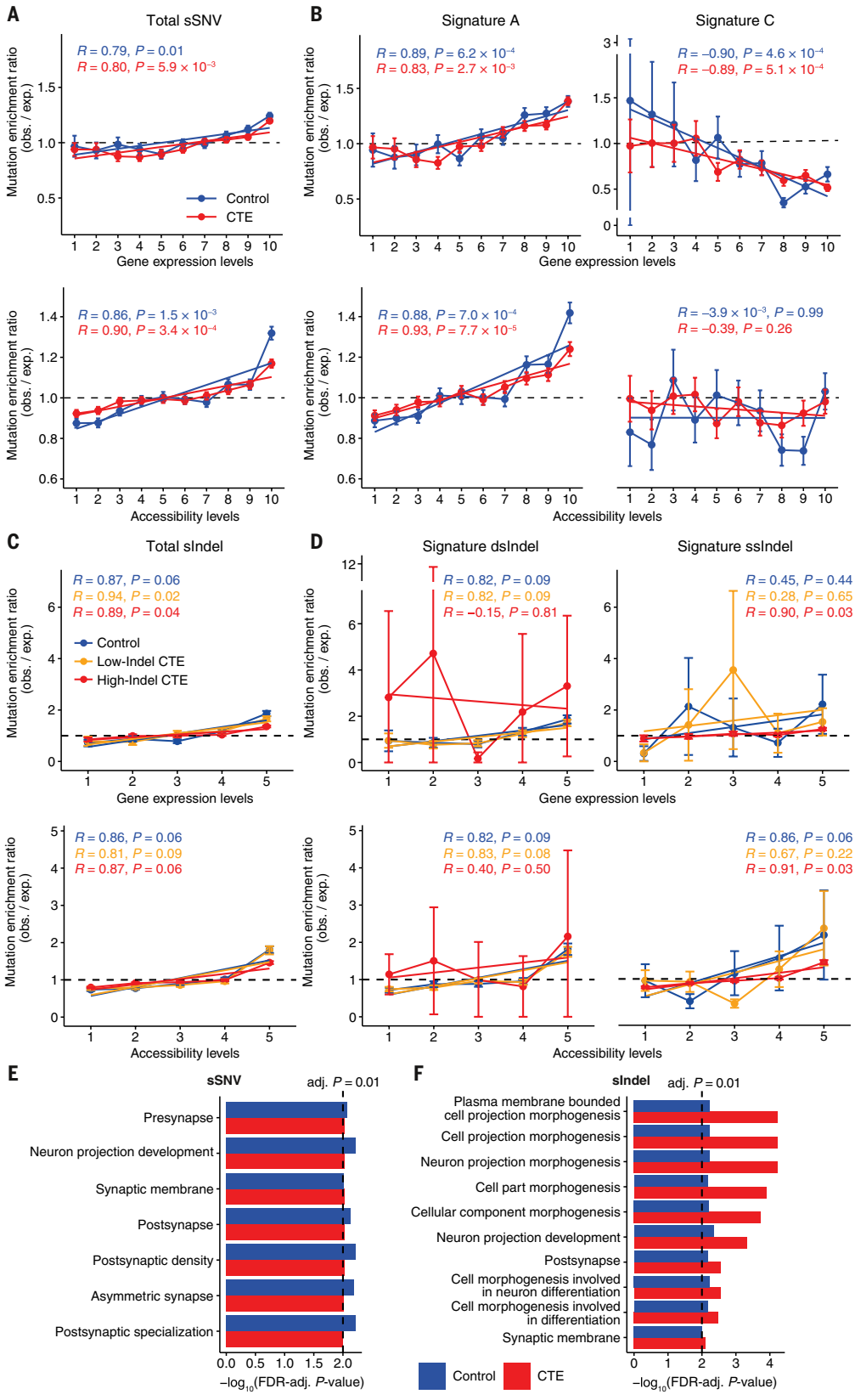


Fig. 5. Enrichment analysis of sSNVs and sIndels in CTE neurons. (**A** to **D**) sSNV [(A) and (B)] and sIndel [(C) and (D)] density as a function of gene expression and chromatin accessibility levels. Neuronal transcriptional profiles were characterized from snRNA-seq data sequenced in this study. Neuronal accessibility profiles were obtained from snATAC-seq data in Ganz *et al*. (*23*). Genes and open chromatin regions are separated into 10 (for sSNV) or 5 (for sIndel) equally sized groups with increasing levels indicating increasing expression or accessibility. Observed density of each expression or accessibility group is obtained by overlapping the original mutation call sets with regions of each group. Expected density is calculated from 1000 permutations of mutation call sets overlapped with regions of each group. Enrichment ratio is calculated by

observed/expected density for each permutation, and mean ratio over 1000 permutations is used to construct the trend line by linear model (error bar indicates standard deviation). Pearson correlation coefficient (*R*) and two-tailed *P* value (*P*) are shown. obs., observed; exp., expected. (**E** and **F**) Gene ontology (GO) analysis of genes where sSNVs (E) and sIndels (F) are located. GO terms with false discovery rate (FDR)–adjusted $P < 0.01$ in both CTE and control are reported.

CTE neurons appeared to have distinct distributions of double- and single-stranded events that may be specific to each individual. Most sSNVs from all CTE individuals were represented on both strands; however, High-Indel CTE individuals present a specific pattern contributed by many features of ID4. ID4 was recently reported to occur when ribonucleotides are misincorporated into DNA and subsequently cleaved by TOP1 instead of RNase H2 (*35*). These double-stranded deletions are thought to arise via an intermediate step of single-stranded >1-bp deletions, which aligns with our observation that ID4-like patterns have a stronger signal in ssIndels. Although the association between RNase H2 deficiency and neurodegenerative disease is unknown, our results raise the possibility that a large quantity of genome-embedded ribonucleotides in CTE and AD neurons might overwhelm the ribonucleotide excision repair pathway, with the TOP1-mediated process leaving large numbers of >1-bp deletions on single DNA strands of which only a small proportion could be efficiently repaired or fixed into double-stranded Indels in postmitotic neurons (*36*). A parallel study (*36*) found a similar sIndel pattern in two other neurodegenerative diseases with distinct pathologies and provided direct experimental evidence of some expected features of TOP1-mediated mutagenesis, suggesting an even wider impact of TOP1 activity in neurodegeneration. It is also worth noting that ID4 may not fully explain the DNA damage in CTE, as some excess sIndels, particularly 2-bp deletions in one repeat unit, are limited in ID4. This suggests that other mutagenic processes may play a role in congruence with or independent of ID4-associated processes. Nonetheless, our analysis highlights that CTE can be viewed as a genomic process in affected neurons, with progressive transcription-related accumulation of sSNVs and especially sIndels—which accumulate to numbers equivalent to >100 years of excess aging—creating the potential for severe dysregulation of the neuronal transcriptome (fig. S17).

Materials and methods

Data reporting

No statistical methods were used to predetermine sample size. Experimenters were not blinded, and experiments were not randomized.

Human tissue samples and CTE case selection

Postmortem frozen human brain tissues were obtained from the UNITE Brain Bank (formerly VA-BU-CLF) at Boston University, the Massachusetts Alzheimer's Disease Research Center (MADRC) at Massachusetts General Hospital, and the NIH Neurobiobank at the University of Maryland Brain and Tissue Bank (UMBTB).

Tissue samples were collected and distributed for research and publication under the protocols approved by the Boston University–Veterans Affairs Institutional Review Board (for BU-VA: S07-02-0087), MassGeneral Brigham Institutional Review Boards (for MADRC: 1999P009556/MGH, expedited waiver category 5), and the University of Maryland Institutional Review Board (for UMBTB: 00042077) and after provision of written authorization and informed consent. Research on these de-identified specimens and data was performed at Boston Children's Hospital with approval from the Committee on Clinical Investigation (S07-02-0087 with waiver of authorization, exempt category 4) and at Brigham and Women's Hospital with approval from the MassGeneral Brigham Institutional Review Board (2019P003790 for secondary use as nonhuman subjects research).

Neurotypical control and AD cases were included from part of previous studies (*17*, *18*, *23*). Neurotypical control cases had no clinical history of neurological disorders, and AD cases were pathologically confirmed Braak stage V-VI. CTE cases were pathologically confirmed stage III-IV (*3*, *42*). Neither AD nor CTE cases had other notable neurodegenerative pathology.

Isolation of individual pyramidal neurons using FANS

Isolation of single neuronal nuclei was performed through fluorescence-activated nuclear sorting (FANS) for the neuronal nuclear transcription factor (NeuN), as described in previous work (*18*, *21*, *43*). In brief, nuclei were prepared from unfixed frozen (at −80°C) human brain tissue in a dounce homogenizer using a chilled tissue lysis buffer [10 mM Tris-HCl, 0.32 M sucrose, 3 mM $Mg(OAc)_2$, 5 mM $CaCl_2$, 0.1 mM EDTA, 1 mM dithiothreitol (DTT), 0.1% Triton X-100, pH 8] on ice. Lysed tissue was then layered on top of a sucrose cushion buffer [1.8 M sucrose 3 mM $Mg(OAc)_2$, 10 mM Tris-HCl, 1 mM DTT, pH 8] and ultracentrifuged for 1 hour at 30,000*g*. Nuclear pellets were resuspended in ice-cold phosphate-buffered saline supplemented with 3 mM $MgCl_2$, filtered, then stained with DAPI and anti-NeuN (RBFOX3) antibody directly conjugated to Alexa Fluor 488 (AF488) (Millipore MAB377X, clone A60, 1:1,250; RRID: AB_2149209). DAPI staining allowed for isolation of single diploid neuronal nuclei apart from debris and doublet droplets. Previous work showed that using flow cytometry (with software BD FACSDiva v.8.0.2) to gate for the nuclei population with the highest NeuN signal as well as the highest forward scatter area (FSC-A) signal produced 99.3% excitatory pyramidal neurons (*18*).

scWGS of pyramidal neurons using PTA

Isolated single nuclei were sorted one nucleus per well into 96-well plates and their genomes were amplified by PTA (*22*). The quasi-linear process of amplification in PTA has been found to reduce amplification artifacts and improve variant calling (*17*). PTA was performed using the ResolveDNA Whole Genome Amplification Kit (BioSkryb Genomics) as described in previous work (*18*). In brief, nuclei were sorted into 3 μl Cell Buffer prechilled on ice and lysed by addition of 3 μl MS Mix, with mixing at 1400 rpm performed after each step. Lysed nuclei were then neutralized with 3 μl SN1 buffer, followed by 3 μl of SDX reagent, a 10-min incubation at room temperature, and 8 μl of reaction mix (containing polymerase) for a 20 μl of total reaction volume. Next, amplification was performed at 30°C for 10 hours, followed by enzyme inactivation at 65°C for 3 min and DNA cleanup using AMPure. The yield was then determined by PicoGreen binding (Quant-iT dsDNA Assay Kit, Thermo Fisher Scientific). Multiplex polymerase chain reaction (PCR) at four random genomic loci was used as quality control (*21*), where samples with positive amplification at all four loci were retained for Illumina sequencing.

Library preparation followed a modified KAPA HyperPlus Library Preparation protocol described in the ResolveDNA EA Whole Genome Amplification protocol. In brief, end repair and A-tailing were performed for 100 to 500 ng amplified DNA input, followed by adapter ligation using the SeqCap Adapter Kit (Roche, 07141548001) and cleanup using AMPure. After on-bead PCR amplification, size selection was carried out using AMPure for libraries with a size of 300 to 600 bp. Quality control was performed on selected libraries using PicoGreen and TapeStation HS DS100 Screen Tape (Agilent PN 5067-5584), followed by paired-end sequencing (150 bp × 2) on Illumina NovaSeq sequencers at 30× coverage (table S2). Data from PTA-amplified neuronal genomes in CTE and RHI were analyzed alongside previously reported data from control (*17*, *23*) and AD neurons (*18*).

scWGS of pyramidal neurons using modified META-CS

The genomes of single neuronal nuclei were amplified by a modified version of META-CS (*19*), a transposase-based whole genome

amplification technique in which each DNA fragment is tagged and barcoded with 16 unique tags [dataset S1 of Xing *et al.* (*19*)], allowing for single-cell, strand-resolved identification.

DNA oligos were ordered from IDT. Each of the 16 META-CS transposons were annealed and assembled into transposomes with Diagenode Tagmentase (Diagenode; C01070010) per manufacturer's protocol and stored at −80°C.

Single neuronal nuclei, isolated as previously described, were sorted one per well into 96-well plates and lysed in 2 μl of 1X Single Cell Lysis Buffer [20 mM Tris, pH 8.0, 20 mM NaCl, 0.15% Triton X-100, 25 mM DTT, 1 mM EDTA, 1.5 mg/ml Thermolabile Proteinase K (TLPK) (NEB, P8111S)] at 30°C for 1 hour, 55°C for 10 min. Single-cell lysates were stored at −20°C if not immediately amplified.

Lysed nuclei were then transposed by the addition of 8 μl transposition mix (5 μl Diagenode 2X Tagmentation buffer (Diagenode; C01019043), 1 μl diluted META-CS transposome, 2 μl H_2O), mixed at 1640 rpm for 1 min, spun down at 1500 rpm, and incubated at 55°C for 15 min. Transposases were removed by the addition of 2 μl 6X stop buffer containing 300 mM NaCl, 45 mM EDTA, 0.01% Triton X-100, and 1 mg/ml TLPK, with mixing and incubation at 37°C for 30 min, 55°C for 10 min.

First-strand tagging was performed by the addition of 13 μl Strand Tagging Mix 1 containing 5 μl Q5 reaction buffer (NEB; B9027S), 5 μl Q5 high GC enhancer (NEB; B9028A), 0.85 μl 100 μM (total) Adp1 primer mix [dataset S1 of Xing *et al.* (*19*)], 0.6 μl 100 mM $MgCl_2$, 0.55 μl water, 0.5 μl 10 mM (each) dNTP mix (Thermo Scientific; R0192), 0.25 μl of 20 mg/ml bovine serum albumin (NEB; B9000S), and 0.25 μl Q5 DNA polymerase (NEB; M0491S), followed with mixing and incubation at 72°C for 3 min, 98°C for 30 s, 62°C for 5 min, 72°C for 1 min. ADP1 primers were removed with 1 μl Thermolabile Exonuclease I (NEB; M0568L), with mixing and incubation at 37°C for 15 min, 65°C for 5 min.

Second-strand tagging was performed by the addition of 4 μl Strand Tagging Mix 2 containing 1 μl Q5 reaction buffer, 1 μl Q5 high GC enhancer, 0.95 μl 100 μM (total) Adp2 primer mix [dataset S1 of Xing *et al.* (*19*)], 0.85 μl water, 0.1 μl 10 mM (each) dNTP mix, and 0.1 μl Q5 DNA polymerase, and incubation at 72°C for 3 min, 98°C for 30 s, 62°C for 5 min, 72°C for 1 min. Adp2 primers were removed by repeating the exonuclease step described above.

Strand tagging products were amplified by the addition of 19 μl PCR mix containing 1 μl NEBNext Multiplex Oligos Universal Primer, 1 μl NEB Index Primers (NEB; E7335S, E7500S, E7710S, E7730S), 4 μl Q5 reaction buffer, 4 μl Q5 high GC enhancer, 0.4 μl 10 mM (each) dNTP mix, 8.4 μl water, and 0.2 μl Q5 DNA polymerase and incubation at 98°C for 20 s, 13 cycles of [98°C for 10 s, 72°C for 2 min], 72°C for 2 min.

Single-cell libraries were pooled together, and then libraries were purified by DNA Clean and Concentrator-5 columns (Zymo; D4013) and amplification efficiency was checked for fragment size and concentration by Agilent TapeStation or Agilent Bioanalyzer. Size selection was performed with Ampure XP beads (Beckman Coulter; A63880), wherein the pooled library was divided into three groups on the basis of fragment size. Medium-size fragments (~300 to 1000 bp) were selected first by the addition of 0.6X beads then by a further addition of 0.25X beads (for a final 0.75X).

snRNA-seq of CTE and control samples

Single-nucleus RNA sequencing (snRNA-seq) was performed on representative tissue samples (control individual UMB1465, prefrontal cortex; CTE individual CTE9130, prefrontal cortex) to assess mutational enrichment with gene expression. Isolation of nuclei was performed as described above with the following modifications: both 0.2 U μl^{-1} Protector RNase inhibitor (Roche RNAINH-RO) and 0.2 U μl^{-1} SUPERase-In RNase inhibitor (Invitrogen) were added to the tissue lysis buffer and to the immunostaining buffer, and $MgCl_2$ was omitted from the immunostaining buffer. For each sample, ~16,000 nuclei were sorted into one well of 96-well plates. snRNA-seq was performed using the 10X Genomics Next GEM Single Cell 3′ GEM Kit v3.1 and Chromium Controller, followed by Illumina sequencing.

Read mapping and BAM file generation for bulk and PTA data

BWA (v.0.7.15) (*44*) was first used to map reads from bulk WGS and PTA scWGS data onto the human reference genome (GRCh37 with decoy) with default parameters. Then, duplicate reads were marked by MarkDuplicates of Picard (v.2.8.0), followed by local Indel realignment and base quality score recalibration using Genome Analysis Toolkit (GATK) (v.3.5) (*45*) to generate BAM files for mutation calling.

Quality measures of single-cell genome amplification

To evaluate the quality of single-cell genome amplification for both PTA and META-CS data, we used a number of measures. First, we calculated the median absolute pairwise differences (MAPD) to quantify the evenness of amplification as described previously (*46*). MAPD score was computed by binning the genome, estimating copy number of each bin, and taking the median of absolute pairwise differences between log2-transformed copy number ratios of adjacent bins. A lower MAPD score indicates more-even amplification of the genome. Further, to account for the variance of the copy number ratio distribution, we calculated the coefficient of variation (CoV) of these absolute pairwise differences by taking the ratio of their standard deviation to their mean. Sequencing depth was estimated using the total number of properly mapped and paired reads (from samtools stats) multiplied by read length and divided by the whole genome length. To account for amplification bias, we estimated the allelic and locus dropout rates using a set of high-quality germline heterozygous SNPs that overlap with common variants from the 1000 Genome Project. Allelic dropout sites have either reference (REF) allele depth or nonreference (ALT) allele depth < 2. Locus dropout sites have total depth < 5. Strand dropout rate was estimated as the square root of allelic dropout rate.

Calling of sSNVs and sIndels from PTA data

We used Single Cell ANalysis 2 (SCAN2, v.1.0) (*17*) to identify sSNVs and sIndels from single-cell PTA data with matched bulk data. First, we generated four cross-sample panels. One panel for 17 control individuals from UMBTB, one panel for 4 RHI individuals, one panel for 15 CTE individuals and 2 control individuals from UNITE Brain Bank (formerly VA-BU-CLF), and one panel for 7 AD individuals. Each panel was configurated by running "scan2 config–analysis makepanel" with the following reference parameters: human reference genome GRCh37 with decoy (--ref), dbSNP v138 common variants (--dbsnp), and 1000 Genomes Phase 3 SHAPEIT2 phasing panel (--shapeit-refpanel). Each panel was then built by running "scan2 makepanel." After panel generation, mutation calling was performed for each individual by running "scan2 config–analysis call_mutations" with the same reference parameters above, each PTA BAM (--sc-bam), matched bulk BAM (--bulk-bam), and the corresponding cross-sample panel (--cross-sample-panel), followed by "scan2 run." For each single cell, SCAN2 generated both mutation calls and genome-wide somatic mutation burden estimations (autosomes only) which were used for the subsequent PTA analyses described below. Two AD cells were excluded from this study either because of failed SCAN2 run (ALZ1647BA9-C) or large discrepancy in burden estimations between SCAN2 and LiRA (1995P_201001E3), the tool originally used to profile these cells (*18*).

Preprocessing and filtering of META-CS data

Our pipeline to preprocess META-CS data and perform sSNV calling shares a core workflow with the previously reported method (*19*) and includes extra steps to further strengthen accuracy and remove false positives, which better accommodates our modified experimental protocol. First, paired-end reads were preprocessed by pre-meta including

identifying transposon barcodes, merging overlapping read ends, and trimming Illumina adapters. Then, reads were mapped to the human reference genome (GRCh37 with decoy) to generate BAM files using BWA-MEM (v.0.7.17) (*47*) and Minimap2 (v.2.12) (*48*) for SNV and BWA-MEM (v.0.7.17) (*47*) for Indel. Of note, because each original DNA fragment was tagged by a pair of transposon barcodes, we split BAM files by barcode pairs before running mutation calling in each barcode pair BAM. This allows us to filter out reads without matching barcodes and ensures that calling was performed on a single-molecule level (for each allele). In addition, given the relatively small number of unique barcodes, there is a chance of barcode collision where different DNA fragments are tagged by the same barcode pair. Therefore, we extracted Tn5 cut sites from each read pair, with the assumption that reads amplified from the same DNA fragment should share the same Tn5 cut sites. Other quality metrics were obtained in the same way as PTA data (described above). Owing to batch effects and imbalanced coverage among pooled cells, cells with an average insert size of <280 or >500, or a standard deviation of insert size >750, were filtered out. In addition, some cells had abnormally high numbers of calls, and we found a significant overlap (~93%) between their prefiltered call sets and the gnomAD (*49*) SNPs of ≥1% population frequency. These cells were removed owing to potential contamination with another individual.

Calling of sSNVs from META-CS data

We generated sSNV candidates by identifying variants that have no ALT allele read from bulk but at least four total ALT reads and at least two ALT reads from each strand for duplex support in the cell. To achieve robust and accurate calling, we filtered out candidate sites that satisfied any of the following criteria: overlapping with the low-quality regions as previously reported (*19*), overlapping with gnomAD (*49*) SNPs of ≥1% population frequency, within 100 bp from another candidate site. Passing variants need to have at least four ALT reads in total with at least two ALT reads from each strand (a4s2) as well as a VAF = 1 at single-molecule level.

Calling of sIndels from META-CS data

We established a new pipeline (https://github.com/gldong/duplex-indel) for sIndel calling from META-CS data by introducing novel modules to expand on the sSNV pipeline (manuscript in preparation). First, as the majority of false positive calls originated from incorrectly merged read pairs during preprocessing, we created a module to tag the merged reads and calculate a genomic window where the merging occurred. Second, we generated an additional BAM file without read merging and filtered out any candidate sites that either overlapped with the read merging window or were not present in the BAM without read merging. In addition, we removed any candidate sites that were adjacent to a germline Indel (located within the flanking region of the germline Indel start site; flanking region is defined by 5 bp or twice of the germline Indel length on both sides, whichever is larger). Other filtering criteria were the same as the ones used for sSNV calling, except that we used gnomAD (*49*) Indels of ≥1% population frequency. Passing variants need to have at least four ALT reads in total with at least two ALT reads from each strand (a4s2) as well as a VAF = 1 at single-molecule level.

Calling of single-stranded variants from META-CS data

Single-stranded SNV and Indel calls were detected in a similar fashion as double-stranded calls described above. We generated candidate sites by identifying variants that have no ALT read from bulk, at least four ALT reads from the variant strand in the cell, and no ALT read from the non-variant strand in the cell. Then, the same filtering strategy except for the ALT allele balance was implemented to generate the final call set. Passing variants need to have at least four ALT reads from the variant strand and at least four REF reads from the nonvariant strand.

Additional filtering of SNV and Indel calls from META-CS

To account for batch effects and technical artifacts, we implemented a set of filters to remove false positives. The same filters were applied to both double- and single-stranded SNVs and Indels. We removed variant sites that meet at least one of the following criteria: (i) variant site is covered by more than one barcode-pair, (ii) barcode-pair covering the variant site has more than one set of Tn5 cut sites, (iii) variant site located within 20 bp of read ends.

Statistical models of somatic mutation burden

We used linear mixed-effects (LME) models from the lme4 (v.1.1-27.1) R package (*50*) to investigate potential associations between somatic mutation burden and other covariates of interest. Somatic mutation burden was modeled as a continuous outcome, covariates of interest including age, clinical status, and cell quality measures were modeled as fixed effects, and individuals were modeled as random effects due to potential correlations between cells from the same individual. *P*-values of fixed effects were obtained from *t* tests with the Satterthwaite approximation using the lmerTest (v.3.1-3) R package (*51*). Each clinical group was tested against controls separately to obtain the corresponding *P* value. To calculate QC-corrected burden, we first modeled each quality measure and age as fixed effects in control neurons to obtain the coefficient of this quality measure's contribution and then subtracted its contribution from total somatic burden of each cell before modeling the corrected somatic burden against age and clinical status to generate QC-corrected burden comparisons.

As disease cells often exhibit large variation in somatic mutation burden (which could also represent biological effects), the various assumptions of the LME model may not always hold. Therefore, to strengthen our statistical testing results, we performed additional nonparametric Wilcoxon tests on excess mutation burdens across clinical groups after adjusting for age. To calculate the age effects on mutation burden, we fitted burdens from control cells in an LME model with age as fixed effects and individuals as random effects, and checked the model assumptions (linearity, homoscedasticity, and normality) (fig. S2). Then we used the coefficient of age to calculate an expected age-contributed mutation burden in cells across clinical groups, before subtracting it from the cell's total mutation burden to obtain excess mutation burden. Finally, we used two-tailed Wilcoxon test, a nonparametric test that does not require those assumptions for LME, to test the difference in excess mutation burdens between pairs of clinical groups.

Association analysis of somatic mutation burden

We used the same LME model described above to evaluate the association between somatic mutation burden and covariates of interest in CTE individuals. For years of playing football, sSNV or sIndel burden was modeled as a continuous outcome, age and the number of years playing football were modeled as fixed effects, and individuals were modeled as random effects. *P* values of fixed effects were obtained from *t* tests with the Satterthwaite approximation using the lmerTest (v.3.1-3) R package (*51*). For *APOE* ε4, we first genotyped all CTE individuals at two SNPs, rs429358 and rs7412, using their bulk WGS data (six possible genotypes are ε2/ε2, ε2/ε3, ε2/ε4, ε3/ε3, ε3/ε4, and ε4/ε4). Then, we used three genetic models of ε4 allele, dominant, additive (with linear penetrance), and recessive, to test the association between *APOE* ε4 genotype and somatic burden. Somatic burden was modeled as a continuous outcome, age and *APOE* ε4 genotype were modeled as fixed effects, and individuals were modeled as random effects. *P* values of fixed effects were obtained from *t* tests with the Satterthwaite approximation using the lmerTest (v.3.1-3) R package (*51*).

Double- and single-stranded signature analysis

Strand-specific signatures were extracted from META-CS call sets and normalized. To better interpret these signatures, we decomposed them

to the COSMIC v3.2 database (https://cancer.sanger.ac.uk/signatures) with matched mutation type using MutationalPatterns (v.3.0.1) R package (*30*).

Because these signatures were derived independently, the double- and single-stranded signatures for each mutation type exhibited collinearity that made them difficult for refitting. Therefore, two intermediate signatures were introduced to serve as a proxy for refitting. For each mutation type, the double- and single-stranded signatures, DS and SS, were combined and normalized, from which two intermediate signatures, IM_1 and IM_2, were extracted using NMF. DS and SS were then refitted to IM_1 and IM_2 so that they could be reconstructed as

$$\widehat{DS} = a_1 \times IM_1 + a_2 \times IM_2,$$
$$\widehat{SS} = b_1 \times IM_1 + b_2 \times IM_2$$

where a_1, a_2, b_1, and b_2 were proportional contributions normalized to IM_1 and IM_2 (i.e., $a_1 + b_1 = 1$ and $a_2 + b_2 = 1$).

To estimate the proportions of double- and single-stranded calls within the PTA call sets, for each mutation type in each cell, PTA calls were fitted to the intermediate signatures

$$\widehat{PTA} = c_1 \times IM_1 + c_2 \times IM_2$$

and the signature contributions were converted to describe DS and SS

$$\widehat{PTA}_{converted} = k_1 \times DS + k_2 \times SS$$

where $k_1 = c_1 \times a_1 + c_2 \times a_2$ and $k_2 = c_1 \times b_1 + c_2 \times b_2$. Note that all signatures and spectra are vectors.

Known and de novo mutational signature analysis

To study the contributions of known and de novo mutational signatures to the PTA call sets, we first categorized the mutation calls from each cell into context groups predefined by COSMIC v3.2 (96 contexts for SNVs and 83 contexts for Indels). For sSNVs, we used MutationalPatterns (*30*) to fit the PTA calls to known signatures A and C reported in Lodato *et al.* (*21*) to get the proportional contribution of each signature to each cell. Then, signature-specific somatic mutation burden was calculated by multiplying the contribution by the genome-wide burden, which is comparable between cells. De novo signatures were extracted using the NMF-based method in MutationalPatterns. The number of signatures to extract was determined with 200 NMF runs based on the Kullback-Leibler divergence (method = "brunet"). PTA calls were fitted to the de novo signatures in the same way as for known signatures. We used the COSMIC v3.2 database to decompose both known and de novo signatures for further interpretation.

Permutation of sSNV and sIndel calls

Permutation sets are crucial for robust enrichment analyses by controlling for potential biases that exist in the original mutation calls and providing powered statistical tests to evaluate the significance of enrichment. For each permutation, the original calls were randomly shuffled across the callable regions of each cell while preserving their chromosome and trinucleotide context. We generated 1000 permutation sets for each mutation type in each cell using SCAN2 by running "scan2 config–analysis permtool" with original mutation calls (--permtool-muts), the human reference genome GRCh37 with decoy (--permtool-bedtools-genome-file), and 1000 permutations (--permtool-n-permutations), followed by "scan2 permtool."

Droplet-based snRNA-seq analysis

Gene count matrices for both control and CTE were acquired by aligning reads to the GRCh37 genome (v.3.0.0) using Cellranger (v.6.1.0) (*52*) with default parameters. Then we used the standard workflow from Seurat (v.4.0.5) (*53*) to process the snRNA-seq data. For the purpose of quality control, we removed genes that were not expressed in <3 cells and cells with (i) <200 genes, (ii) <500 counts or >15,000 counts, or (iii) mitochondrial gene percentage of >5%. After applying these filters, we obtained 4619 cells for the control dataset and 4080 cells for the CTE dataset. Next, the data were normalized using the function "LogNormalize" with "scale.factor" = 10000 and scaled to a maximum value of 10. We then performed dimension reduction (principal components analysis, t-distributed stochastic neighbor embedding, and uniform manifold approximation and projection) and cell clustering using Louvain (*54*). To identify the cell types that were captured in the snRNA-seq data, we selected marker genes for each cluster by statistical tests and assigned cell-type labels according to previously reported marker lists (*55*). Through visual evaluation, cells with the same identity were clustered together.

Enrichment analysis with snRNA-seq and snATAC-seq data

We performed enrichment analysis of somatic mutations in transcribed gene or open chromatin regions using our in-house snRNA-seq data and previously reported snATAC-seq data generated from matching control samples (*23*). We extracted the gene expression and chromatin accessibility profiles of neurons annotated in snRNA-seq and snATAC-seq data to match the PTA data.

For gene expression enrichment analysis, we used the expression profile of control for mutation calls from control cells, and the expression profile of CTE for mutation calls from CTE cells. First, genes were ranked on the basis of their expression and divided into equal-sized groups (10 groups for sSNVs, 5 groups for sIndels). Then we intersected both original and permuted mutation calls with gene regions of each expression group. Gene regions were annotated by ANNOVAR (*56*) using the database GRCh37 refGene (*57*). Mutation calls that overlapped with multiple genes were removed. A ratio of observed to expected number of calls (i.e., enrichment ratio) was calculated for each permutation round where the observed number came from the original call set and the expected number came from the permuted call set. Enrichment was reported as the mean and standard deviation of the 1000 ratios. To test enrichment differences between transcribed and nontranscribed genes, we defined genes with an average expression > 0 as transcribed and the remaining genes as nontranscribed in either control or CTE sample after QC and normalization described above. We counted mutation numbers in transcribed versus nontranscribed genes for each permutation and divided the observed counts over permuted counts to calculate an enrichment ratio for each transcriptional status in each clinical group. Then, a two-tailed Wilcoxon test was performed between transcribed and nontranscribed enrichment ratios in each group.

For chromatin accessibility enrichment analysis, we used the processed BED file of excitatory neurons (*23*) where the genome was first separated into nonoverlapping 1000-bp bins and snATAC-seq fragments were mapped to the bins. Similar to gene expression enrichment, the bins were ranked on the basis of their coverage and divided into equal-sized groups (10 groups for sSNVs, 5 groups for sIndels). Then we intersected both original and permuted mutation calls with genomic regions of each accessibility group. Enrichment ratios were calculated in the same way as above.

In addition to total somatic mutation enrichment, we performed signature-specific enrichment using the same NMF-based method for signature analysis above. For both original and permuted calls that overlapped with each expression or accessibility group, we fit the sSNV calls to signatures A and C and the sIndel calls to the double- and single-stranded signatures from META-CS. The observed/expected ratios were calculated as above.

Gene Ontology analysis

Gene Ontology enrichment analysis was performed on genes that overlapped with somatic mutations using GOseq (v.1.42.0) (*58*) after controlling for gene length bias. A null distribution was generated by

Wallenius approximation, and GO categories with less than 10 hits or more than 1000 genes were filtered out. P values of over- and under-represented GO categories were adjusted for multiple testing using FDR. All GO categories with adjusted $P < 0.05$ in both CTE and neurotypical controls were reported in table S7.

REFERENCES AND NOTES

1. A. C. McKee *et al.*, The first NINDS/NIBIB consensus meeting to define neuropathological criteria for the diagnosis of chronic traumatic encephalopathy. *Acta Neuropathol.* **131**, 75–86 (2016). doi: 10.1007/s00401-015-1515-z; pmid: 26667418
2. A. C. McKee, T. D. Stein, P. T. Kiernan, V. E. Alvarez, The neuropathology of chronic traumatic encephalopathy. *Brain Pathol.* **25**, 350–364 (2015). doi: 10.1111/bpa.12248; pmid: 25904048
3. K. F. Bieniek *et al.*, The second NINDS/NIBIB consensus meeting to define neuropathological criteria for the diagnosis of chronic traumatic encephalopathy. *J. Neuropathol. Exp. Neurol.* **80**, 210–219 (2021). doi: 10.1093/jnen/nlab001; pmid: 33611507
4. R. C. Turner, B. P. Lucke-Wold, M. J. Robson, J. M. Lee, J. E. Bailes, Alzheimer's disease and chronic traumatic encephalopathy: Distinct but possibly overlapping disease entities. *Brain Inj.* **30**, 1279–1292 (2016). doi: 10.1080/02699052.2016.1193631; pmid: 27715315
5. R. A. Stern *et al.*, Clinical presentation of chronic traumatic encephalopathy. *Neurology* **81**, 1122–1129 (2013). doi: 10.1212/WNL.0b013e3182a55f7f; pmid: 23966253
6. A. C. McKee *et al.*, The spectrum of disease in chronic traumatic encephalopathy. *Brain* **136**, 43–64 (2013). doi: 10.1093/brain/aws307; pmid: 23208308
7. B. Falcon *et al.*, Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420–423 (2019). doi: 10.1038/s41586-019-1026-5; pmid: 30894745
8. Y. Shi *et al.*, Structure-based classification of tauopathies. *Nature* **598**, 359–363 (2021). doi: 10.1038/s41586-021-03911-7; pmid: 34588692
9. P. H. Montenigro, C. Bernick, R. C. Cantu, Clinical features of repetitive traumatic brain injury and chronic traumatic encephalopathy. *Brain Pathol.* **25**, 304–317 (2015). doi: 10.1111/bpa.12250; pmid: 25904046
10. B. Vogelstein *et al.*, Cancer genome landscapes. *Science* **339**, 1546–1558 (2013). doi: 10.1126/science.1235122; pmid: 23539594
11. M. R. Stratton, P. J. Campbell, P. A. Futreal, The cancer genome. *Nature* **458**, 719–724 (2009). doi: 10.1038/nature07943; pmid: 19360079
12. L. A. Garraway, E. S. Lander, Lessons from the cancer genome. *Cell* **153**, 17–37 (2013). doi: 10.1016/j.cell.2013.03.002; pmid: 23540688
13. I. Martincorena *et al.*, Somatic mutant clones colonize the human esophagus with age. *Science* **362**, 911–917 (2018). doi: 10.1126/science.aau3879; pmid: 30337457
14. I. Martincorena *et al.*, High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* **348**, 880–886 (2015). doi: 10.1126/science.aaa6806; pmid: 25999502
15. G. Genovese *et al.*, Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014). doi: 10.1056/NEJMoa1409405; pmid: 25426838
16. F. Abascal *et al.*, Somatic mutation landscapes at single-molecule resolution. *Nature* **593**, 405–410 (2021). doi: 10.1038/s41586-021-03477-4; pmid: 33911282
17. L. J. Luquette *et al.*, Single-cell genome sequencing of human neurons identifies somatic point mutation and indel enrichment in regulatory elements. *Nat. Genet.* **54**, 1564–1571 (2022). doi: 10.1038/s41588-022-01180-2; pmid: 36163278
18. M. B. Miller *et al.*, Somatic genomic changes in single Alzheimer's disease neurons. *Nature* **604**, 714–722 (2022). doi: 10.1038/s41586-022-04640-1; pmid: 35444284
19. D. Xing, L. Tan, C. H. Chang, H. Li, X. S. Xie, Accurate SNV detection in single cells by transposon-based whole-genome amplification of complementary strands. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2013106118 (2021). doi: 10.1073/pnas.2013106118; pmid: 33593904
20. M. H. Liu *et al.*, DNA mismatch and damage patterns revealed by single-molecule sequencing. *Nature* **630**, 752–761 (2024). doi: 10.1038/s41586-024-07532-8; pmid: 38867045
21. M. A. Lodato *et al.*, Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* **359**, 555–559 (2018). doi: 10.1126/science.aao4426; pmid: 29217584
22. V. Gonzalez-Pena *et al.*, Accurate genomic variant detection in single cells with primary template-directed amplification. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2024176118 (2021). doi: 10.1073/pnas.2024176118; pmid: 34099548
23. J. Ganz *et al.*, Contrasting somatic mutation patterns in aging human neurons and oligodendrocytes. *Cell* **187**, 1955–1970.e23 (2024). doi: 10.1016/j.cell.2024.02.025; pmid: 38503282
24. J. Mez *et al.*, Duration of American football play and chronic traumatic encephalopathy. *Ann. Neurol.* **87**, 116–131 (2020). doi: 10.1002/ana.25611; pmid: 31589352
25. M. L. Alosco *et al.*, Characterizing tau deposition in chronic traumatic encephalopathy (CTE): Utility of the McKee CTE staging scheme. *Acta Neuropathol.* **140**, 495–512 (2020). doi: 10.1007/s00401-020-02197-9; pmid: 32778942
26. A. C. McKee *et al.*, Neuropathologic and clinical findings in young contact sport athletes exposed to repetitive head impacts. *JAMA Neurol.* **80**, 1037–1050 (2023). doi: 10.1001/jamaneurol.2023.2907; pmid: 37639244
27. Z. Sondka *et al.*, COSMIC: A curated database of somatic variants and clinical data for cancer. *Nucleic Acids Res.* **52**, D1210–D1217 (2024). doi: 10.1093/nar/gkad986; pmid: 38183204
28. M. Jager *et al.*, Deficiency of nucleotide excision repair is associated with mutational signature observed in cancer. *Genome Res.* **29**, 1067–1077 (2019). doi: 10.1101/gr.246223.118; pmid: 31221724
29. A. R. Poetsch, The genomics of oxidative DNA damage, repair, and resulting mutagenesis. *Comput. Struct. Biotechnol. J.* **18**, 207–219 (2020). doi: 10.1016/j.csbj.2019.12.013; pmid: 31993111
30. F. Blokzijl, R. Janssen, R. van Boxtel, E. Cuppen, MutationalPatterns: Comprehensive genome-wide analysis of mutational processes. *Genome Med.* **10**, 33 (2018). doi: 10.1186/s13073-018-0539-0; pmid: 29695279
31. I. T. Agaku *et al.*, Combustible and smokeless tobacco use among high school athletes—United States, 2001–2013. *MMWR Morb. Mortal. Wkly. Rep.* **64**, 935–939 (2015). doi: 10.15585/mmwr.mm6434a2; pmid: 26334565
32. F. Chagué *et al.*, Smokeless tobacco, sport and the heart. *Arch. Cardiovasc. Dis.* **108**, 75–83 (2015). doi: 10.1016/j.acvd.2014.10.003; pmid: 25497687
33. A. Serrano-Pozo, S. Das, B. T. Hyman, APOE and Alzheimer's disease: Advances in genetics, pathophysiology, and therapeutic approaches. *Lancet Neurol.* **20**, 68–80 (2021). doi: 10.1016/S1474-4422(20)30412-9; pmid: 33340485
34. K. Atherton *et al.*, Association of APOE genotypes and chronic traumatic encephalopathy. *JAMA Neurol.* **79**, 787–796 (2022). doi: 10.1001/jamaneurol.2022.1634; pmid: 35759276
35. M. A. M. Reijns *et al.*, Signatures of TOP1 transcription-associated mutagenesis in cancer and germline. *Nature* **602**, 623–631 (2022). doi: 10.1038/s41586-022-04403-y; pmid: 35140396
36. Z. Zhou *et al.*, Recurrent patterns of widespread neuronal genomic damage shared by major neurodegenerative disorders. bioRxiv 2025.03.03.641186 [Preprint] (2025); https://doi.org/10.1101/2025.03.03.641186.
37. R. Madabhushi *et al.*, Activity-induced DNA breaks govern the expression of neuronal early-response genes. *Cell* **161**, 1592–1605 (2015). doi: 10.1016/j.cell.2015.05.032; pmid: 26052046
38. W. Wu *et al.*, Neuronal enhancers are hotspots for DNA single-strand break repair. *Nature* **593**, 440–444 (2021). doi: 10.1038/s41586-021-03468-5; pmid: 33767446
39. B. Jin *et al.*, Neurons accumulate disease-specific somatic genomic changes across tau pathologic states in Alzheimer's disease. bioRxiv 2025.05.26.656152 [Preprint] (2025); https://doi.org/10.1101/2025.05.26.656152.
40. A. C. McKee *et al.*, Chronic traumatic encephalopathy in athletes: Progressive tauopathy after repetitive head injury. *J. Neuropathol. Exp. Neurol.* **68**, 709–735 (2009). doi: 10.1097/NEN.0b013e3181a9d503; pmid: 19535999
41. J. D. Cherry, K. J. Babcock, L. E. Goldstein, Repetitive head trauma induces chronic traumatic encephalopathy by multiple mechanisms. *Semin. Neurol.* **40**, 430–438 (2020). doi: 10.1055/s-0040-1713620; pmid: 32674181
42. A. Fesharaki-Zadeh, Chronic traumatic encephalopathy: A brief overview. *Front. Neurol.* **10**, 713 (2019). doi: 10.3389/fneur.2019.00713; pmid: 31333567
43. G. D. Evrony *et al.*, Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* **151**, 483–496 (2012). doi: 10.1016/j.cell.2012.09.035; pmid: 23101622
44. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). doi: 10.1093/bioinformatics/btp324; pmid: 19451168
45. A. McKenna *et al.*, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010). doi: 10.1101/gr.107524.110; pmid: 20644199
46. X. Cai *et al.*, Single-cell, genome-wide sequencing identifies clonal somatic copy-number variation in the human brain. *Cell Rep.* **8**, 1280–1289 (2014). doi: 10.1016/j.celrep.2014.07.043; pmid: 25159146
47. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 [q-bio.GN] (2013).
48. H. Li, Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018). doi: 10.1093/bioinformatics/bty191; pmid: 29750242
49. K. J. Karczewski *et al.*, The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020). doi: 10.1038/s41586-020-2308-7; pmid: 32461654
50. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015). doi: 10.18637/jss.v067.i01
51. A. Kuznetsova, P. B. Brockhoff, R. H. B. Christensen, lmerTest package: Tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017). doi: 10.18637/jss.v082.i13
52. G. X. Zheng *et al.*, Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017). doi: 10.1038/ncomms14049; pmid: 28091601
53. Y. Hao *et al.*, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021). doi: 10.1016/j.cell.2021.04.048; pmid: 34062119

54. V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008). doi: 10.1088/1742-5468/2008/10/P10008
55. H. Mathys *et al.*, Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019). doi: 10.1038/s41586-019-1195-2; pmid: 31042697
56. K. Wang, M. Li, H. Hakonarson, ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010). doi: 10.1093/nar/gkq603; pmid: 20601685
57. N. A. O'Leary *et al.*, Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745 (2016). doi: 10.1093/nar/gkv1189; pmid: 26553804
58. M. D. Young, M. J. Wakefield, G. K. Smyth, A. Oshlack, Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* **11**, R14 (2010). doi: 10.1186/gb-2010-11-2-r14; pmid: 20132535
59. G. Dong, gldong/CTE_manuscript_code: v1.0, Zenodo (2025); https://doi.org/10.5281/zenodo.15300743.

ACKNOWLEDGMENTS

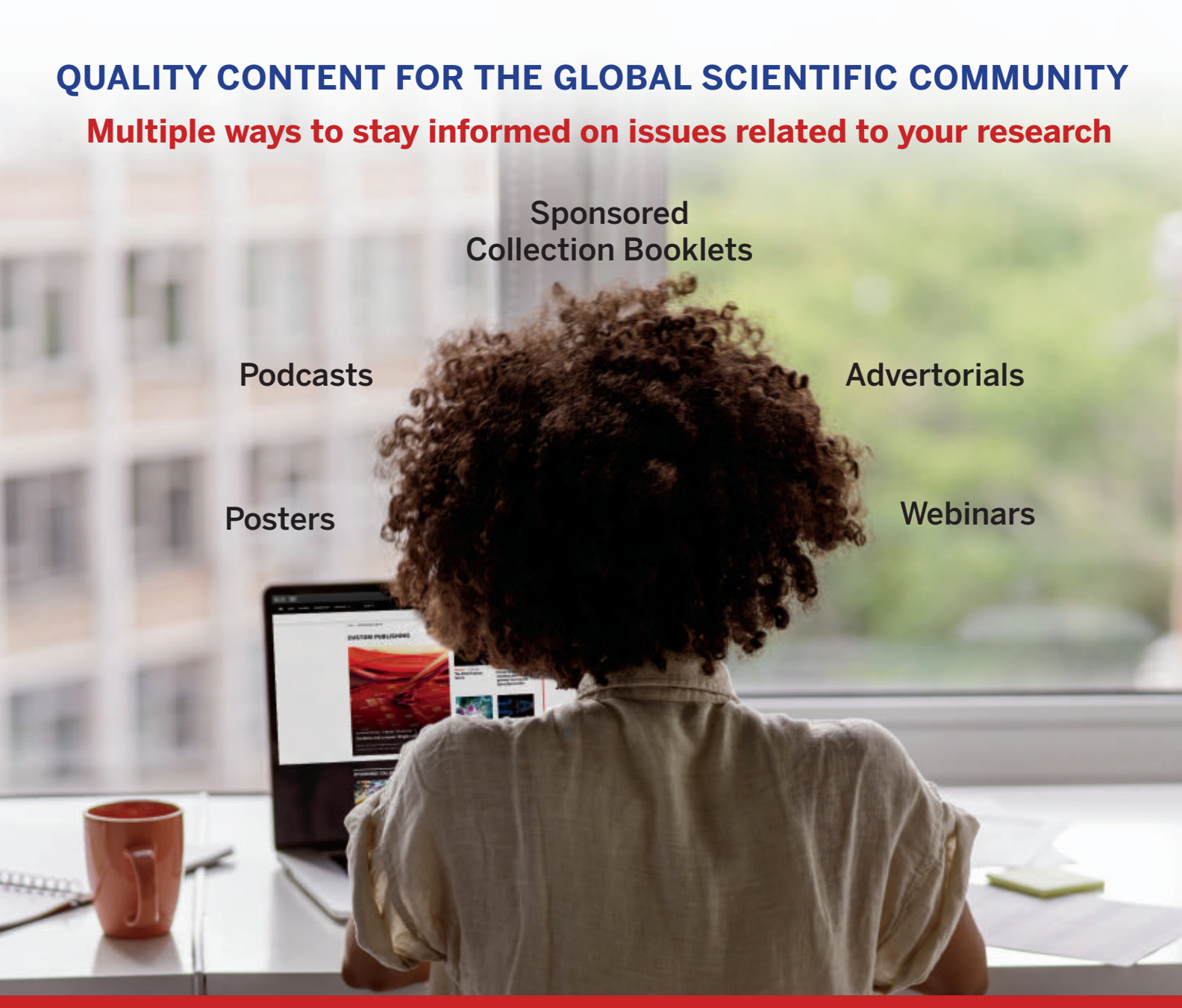
We thank R. Mathieu and T. Berisha from the Boston Children's Hospital Flow Cytometry Core, R. Krishnan from the Brigham and Women's Hospital Center for Neurologic Diseases Flow Cytometry Core Facility, and R. S. Hill and J. E. Neil for their administrative help. We thank the donors of postmortem tissues for their invaluable contributions to the advancement of science. **Funding:** National Institutes of Health (NIH) R56 AG079857 to A.Y.H., C.A.W., and E.A.L.; NIH R01 AG088082 to A.Y.H.; Alzheimer's Association Research Fellowship to A.Y.H.; NIH K08 AG065502 to M.B.M.; NIH DP2 AG086138 to M.B.M.; NIH R01 AG082346 to M.B.M.; Alzheimer's Disease Research program of the BrightFocus Foundation A20201292F to M.B.M.; Doris Duke Charitable Foundation Clinical Scientist Development Award 2021183 to M.B.M.; Suh Kyungbae Foundation to E.A.L.; NIH DP2 AG072437 to E.A.L.; NIH R01 NS032457-20S1 to C.A.W.; NIH R01 AG070921 and AG078929 to C.A.W. and E.A.L.; Templeton Foundation to C.A.W.; Allen Discovery Center program, a Paul G. Allen Frontiers Group advised program of the Paul G. Allen Family Foundation to C.A.W. and E.A.L.; Howard Hughes Medical Institute to C.A.W.; National Institute of Neurological Disorders and Stroke U01NS086659, U01NS093334, U54NS115266, R01NS078337, R56NS078337, and K23NS102399 to A.C.M.; National Institute of Aging P30AG13846 to A.C.M.; Department of Veterans Affairs I01 CX001135 to A.C.M.; National Operating Committee on Standards for Athletic Equipment to A.C.M.; Nick and Lynn Buoniconti Foundation to A.C.M.; Concussion Legacy Foundation to A.C.M.; Andlinger Foundation to A.C.M.; WWE to A.C.M.; NFL to A.C.M. **Author contributions:** M.B.M., A.C.M., and C.A.W. conceived of the study, which was further developed by C.C.M. C.C.M. and M.B.M. performed tissue dissection, single-nuclei sorting, PTA genome amplification, and oversaw sequencing, with assistance from S.M.N., S.L.K., and G.A.M. K.S.-M.B., S.M.N., and S.P.W. performed META-CS genome amplification, with assistance from C.C.M. and M.B.M. and advice and interpretation from D.D.S. G.D. performed data analysis, with assistance from S.M. and J.K. G.D. developed and applied the software for mutation calling in META-CS data. A.C.M., M.U., J.D.C., and E.S. provided brain samples and clinicopathological data and assisted with data interpretation. G.D., C.C.M., S.M., and K.S.-M.B. prepared display items. G.D. and C.C.M. wrote the initial draft of the manuscript, with assistance from M.B.M., A.Y.H., S.M., and K.S.-M.B. M.B.M., A.Y.H., and E.A.L. oversaw experiments and analyses. C.A.W., E.A.L., M.B.M., and A.Y.H. jointly supervised the study and edited the manuscript. **Competing interests:** C.A.W. is a paid consultant (cash, no equity) to Third Rock Ventures and Flagship Pioneering and is on the clinical advisory board (cash and equity) of Maze Therapeutics. E.A.L. is on the scientific advisory board (cash, no equity) of Inocras. No research support was received from the four aforementioned companies. These companies did not fund and had no role in the conception or execution of this research project. G.D., C.C.M., A.Y.H., M.B.M., E.A.L., and C.A.W. have filed a preliminary patent application around findings reported in this paper. All other authors have no competing interests to declare. **Data and materials availability:** Sequencing data generated in this study have been deposited in the National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS) under accession number NG00183, with controlled use conditions set by human privacy regulations. Previously published PTA data for neurotypical controls can be downloaded from the database of Genotypes and Phenotypes (dbGaP) (accession number phs001485.v3.p1), and PTA data for AD can be downloaded from NIAGADS (accession number NG00121). All code used in this study is publicly available in Zenodo (*59*). **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the Author Accepted Manuscript (AAM) of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adu1351
Figs. S1 to S17; Tables S1 to S7; MDAR Reproducibility Checklist
Submitted 24 October 2024; accepted 15 September 2025
10.1126/science.adu1351

QUALITY CONTENT FOR THE GLOBAL SCIENTIFIC COMMUNITY

Multiple ways to stay informed on issues related to your research



Science
AAAS

Scan the code and start exploring the latest advances in science and technology innovation!

Science.org/custom-publishing

Brought to you by the *Science*/AAAS Custom Publishing Office.

Posters **Podcasts** **Sponsored Collection Booklets** **Advertorials** **Webinars**

RESEARCH ARTICLE SUMMARY

CATALYSIS

Conversion of syngas into olefins with high hydrogen atom economy

Chang Gao†, Wenlong Song†, Huiqiu Wang†, Xiao Chen, Chaojie Cui*, Wangshu Hao, Ning Yan, Yuan Yang, Shenglong Yang, Hao Lv, Mingyu Ma, Xinli Lian, Ruixia Zhang, Weizhong Qian*

Full article and list of author affiliations: https://doi.org/10.1126/science.aea0774

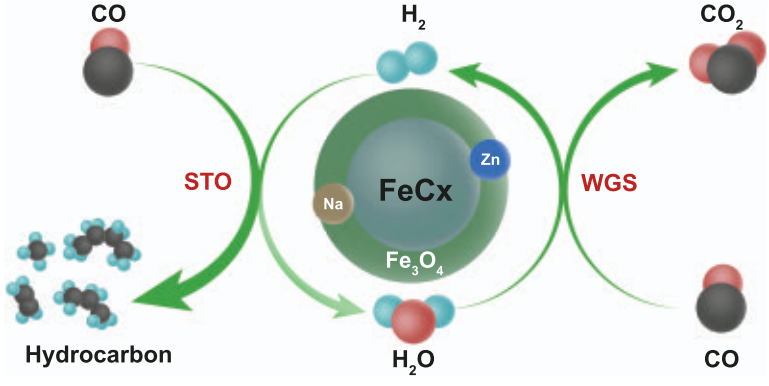
INTRODUCTION: The synthesis of olefins from syngas through methanol has long been hindered by a low hydrogen atom economy (HAE), primarily due to the formation of water by-products and the requirement for high H_2/CO ratios. This limitation not only increases production costs but also raises environmental concerns, prompting the need for innovative catalytic solutions in the emerging technology of direct syngas to olefins.

RATIONALE: To address this challenge, we developed a sodium-modified FeCx@Fe_3O_4 core-shell catalyst that integrates the water-gas shift (WGS) and syngas-to-olefins (STO) reactions. By coupling these two processes, the catalyst enables in situ conversion of water generated during STO into hydrogen through the WGS reaction, thereby reducing the dependence on external hydrogen sources and improving HAE.

RESULTS: Under the conditions of 623 K and 2 MPa with a low H_2/CO ratio feedstock, the catalyst achieved outstanding performance: a CO conversion rate of ~95%, an olefin selectivity exceeding 75% (based on hydrocarbon products), and a hydrocarbon yield of 33 wt % (based on feedstock). The catalyst performance was stable over 500 hours of testing. Its HAE reached ~66 to 86% within an H_2/CO ratio range of 1 to 3, far surpassing the ~43 to 47% of the traditional WGS–methanol synthesis–methanol to olefins (MTO) route (at an H_2/CO ratio of ~2 to 2.05). Isotope-tracing and WGS-blocking experiments validated the coupling mechanism, and the contribution of the WGS reaction was quantitatively determined.

CONCLUSION: This study represents a substantial breakthrough in enhancing HAE for syngas conversion. The new WGS-STO coupling route of the developed catalyst reduces steam consumption, wastewater generation, and CO_2 emissions, with a 46% reduction in complete environmental factor compared with the WGS-MTO route. It also provides a sustainable alternative to existing olefin production technologies, offering great potential for the resource-efficient transformation of the olefin industry and contributing to the achievement of carbon neutrality goals. □

*Corresponding author. Email: qianwz@tsinghua.edu.cn (W.Q.); cuicj06@tsinghua.edu.cn (C.C.) The list of author affiliations is available in the full article online. †These authors contributed equally to this work. **Cite this article as C. Gao *et al.*, *Science* 390, eaea0774 (2025). DOI: 10.1126/science.aea0774**



Improvement of hydrogen atom economy and olefin yield under the WGS-STO coupling effect.

CATALYSIS

Conversion of syngas into olefins with high hydrogen atom economy

Chang Gao[1]†, Wenlong Song[1]†, Huiqiu Wang[1]†, Xiao Chen[1], Chaojie Cui[1,2]*, Wangshu Hao[2], Ning Yan[2], Yuan Yang[2], Shenglong Yang[2], Hao Lv[2], Mingyu Ma[1], Xinli Lian[2], Ruixia Zhang[2], Weizhong Qian[1,2]*

In synthesizing olefins from syngas, low hydrogen atom economy (HAE), the fraction of reactant hydrogen in the hydrocarbon product, arises from hydrogen loss in water by-product. We report a sodium-modified FeC_x@Fe_3O_4 core-shell catalyst coupling water-gas shift (WGS) with syngas-to-olefins (STO) to convert water into hydrogen in situ. HAE reaches about 66 to 83%, exceeding that of methanol-to-olefins (MTO, 50% upper limit). The approximately 95% carbon monoxide conversion and >75% olefin selectivity were simultaneously obtained. The coupling effect was validated by isotope tracing with deuterium oxide and blocking the WGS pathway, and the contribution of WGS was quantitatively evaluated. These results, using lower hydrogen to carbon monoxide ratios, implied that reducing steam consumption in the WGS reaction and reducing the overall output of carbon dioxide and wastewater enabled a sustainable STO process for potential industrialization.

Direct synthesis of olefins from syngas can enable the production of value-added chemicals from versatile resources that substitute for petroleum oil, such as coal, biomass, or natural gas (*1*, *2*). The main challenge of this process is the simultaneous achievement of high conversion of CO, high selectivity to olefins, and low consumption of H_2 (*1*–*4*). Generally, the synthesis of olefins needs a syngas feedstock with an H_2/CO stoichiometric ratio of 2:1. Thus, the water-gas shift (WGS) reaction constitutes a critical step to convert the crude coal-based syngas (with H_2/CO ratio <0.8:1 in most cases) into the H_2-rich syngas in the industry (*5*–*7*).

CO_2 generated during the WGS stage comprises ~60% of the total CO_2 emissions in the industrial coal gasification–WGS–methanol synthesis–methanol to olefins (MTO) pathway (*8*), underscoring the environmental footprint of this process. However, the water generated in the syngas conversion means that hydrogen is not fully used to produce high-value hydrocarbons. The more water produced, the lower the value of hydrogen atom economy (HAE) of the process will be, considering H_2 is the most expensive raw material in this process (*9*).

We describe a Na/FeC_x@Fe_3O_4 core-shell nanoparticle (NP) catalyst that simultaneously achieved improved short-chain olefin yield, high HAE, and low water yield at a low H_2/CO ratio of 1 to 1.5. The intimate contact between the FeC_x (Fe_5C_2 and Fe_3C intergrowth) core and the polycrystalline Fe_3O_4 shell enabled coupling of WGS and syngas-to-olefin (STO), and the nanochannels of the porous shell provided pathways for reactant and product diffusion. By artificially blocking the path, we could quantitatively evaluate the contribution of WGS under different H_2/CO feedstocks. Under optimized conditions, we achieved CO conversion >95%, selectivity to olefins >75% (hydrocarbon base), and hydrocarbon yields of ~33 to 35% by weight (feedstock base). Process analysis suggested that the reaction coupling effect decreased the CO_2 emission at the WGS-STO operation unit, steam consumption at the WGS unit, and wastewater production at the STO unit, which is beneficial for the sustainable development of syngas technology.

Core-shell NP structure and HAE performance

The catalyst precursor was Na-$ZnFe_2O_4$ spinel NPs with an average diameter of 7.6 nm (fig. S1A). During the syngas conversion reaction, we observed the gradual evolution of the distribution of Zn and Fe (fig. S1, B to D), as well as the growth of spinel NPs and the formation of a core-shell sphere-like structure. We proposed that the coalescence and phase precipitation of $ZnFe_2O_4$ NPs resulted from the decrease in melting point after adding Na salts in rational amounts. As the reaction proceeded, atomic annular dark field scanning transmission electron microscopy (ADF-STEM) and energy-dispersive spectroscopy (EDS) mapping images (Fig. 1, A to F) revealed that the spinel NPs decomposed and subsequently transformed into Na-doped ZnO and FeC_x@porous-Fe_3O_4 core-shell NPs (fig. S1C).

The phase composition of the spent catalyst was further confirmed by x-ray diffraction (XRD) analysis (fig. S2), corroborating the structural evolution observed by microscopy. The FeC_x core consisted of Fe_5C_2 and Fe_3C phases (Fig. 1, B, E, and F), both of which exhibited well-documented catalytic activity for STO (*10*, *11*). The porous polycrystalline shell consisted of Fe_3O_4 nanocrystals (Fig. 1E) that are an active phase for WGS reaction. The Na species were distributed not only on the external surface but also in the core, where Na served as a key additive to increase the selectivity of olefins, as discussed below.

This catalyst exhibited high performance when varying the H_2/CO ratios in the feedstock (Fig. 1G). CO conversion was 66% under an H_2/CO ratio of 1, but increased to between 90 and 96% under H_2/CO ratios of 1.25 to 3. The total hydrocarbon selectivity remained at a high level under H_2/CO ratios of 1.25 to 3. Substantial changes were observed in the selectivity of H_2O and CO_2 under different H_2/CO ratios, and higher selectivity of CO_2 was always accompanied by lower selectivity of H_2O.

The lowest output of H_2O was associated with the highest HAE value for hydrocarbon production, which was 88.4% for the feedstock H_2/CO ratio of 1 (Fig. 1H). Water output increased with the increase of H_2/CO ratio in the feedstock but did not give an appreciable increase in hydrocarbon selectivity. As a result, HAE values gradually decreased with the increase of H_2/CO ratio. However, the absolute HAE value in the present work always remained in a very high range (56.7 to 88.4%) (Fig. 1H), well above the upper limit (50%) of MTO and the practical value (41 to 45%) of the commercial MTO route, which had to use the feedstock of H_2/CO ratios of 2.00 to 2.05 (table S1).

In addition, the hydrogen element entering into hydrocarbons mostly contributed to the olefin formation. The feedstock with H_2/CO ratios of 1.25 to 1.5 gave a very similar olefin distribution to that of H_2/CO of 2 (Fig. 1I and fig. S3). The hydrocarbon product profiles were mainly composed of C_1 to C_6 species, whereas the selectivity of C_{7+} species was always lower than the values predicted by the Anderson-Schulz-Flory model (*12*, *13*) (fig. S4 and table S2). In most cases, the ratio of olefin to paraffin (O/P) for the same carbon number ranged from 6 to 7. The highest O/P ratio reached up to ~10 to 12 for C_3 species (table S3). Such high olefin selectivity resulted from the addition of Na species or K species (fig. S5), which were effective in tailoring the local electron state of the catalyst and favored the desorption of olefins from the catalyst. The removal of Na salts by washing with water resulted in the rapid decrease of CO conversion and a decrease of the O/P ratio (from 7.5 to 3.5) within 18 hours (fig. S6).

Finally, a long-term test was performed with an H_2/CO ratio of 2 (Fig. 1J). CO conversion was always >95%, and selectivity of C_2 to C_6 olefins was always at 75% (hydrocarbon base). The excellent stability of the catalyst implied that the water produced (existing as steam) did not degrade the core-shell catalyst with Na additives. High HAE of STO can be maintained for 500 hours. We compared the performance

[1]Department of Chemical Engineering, Tsinghua University, Beijing, China. [2]Ordos Laboratory, Inner Mongolia, China. *Corresponding author. Email: qianwz@tsinghua.edu.cn (W.Q.); cuicj06@tsinghua.edu.cn (C.C.) †These authors contributed equally to this work.

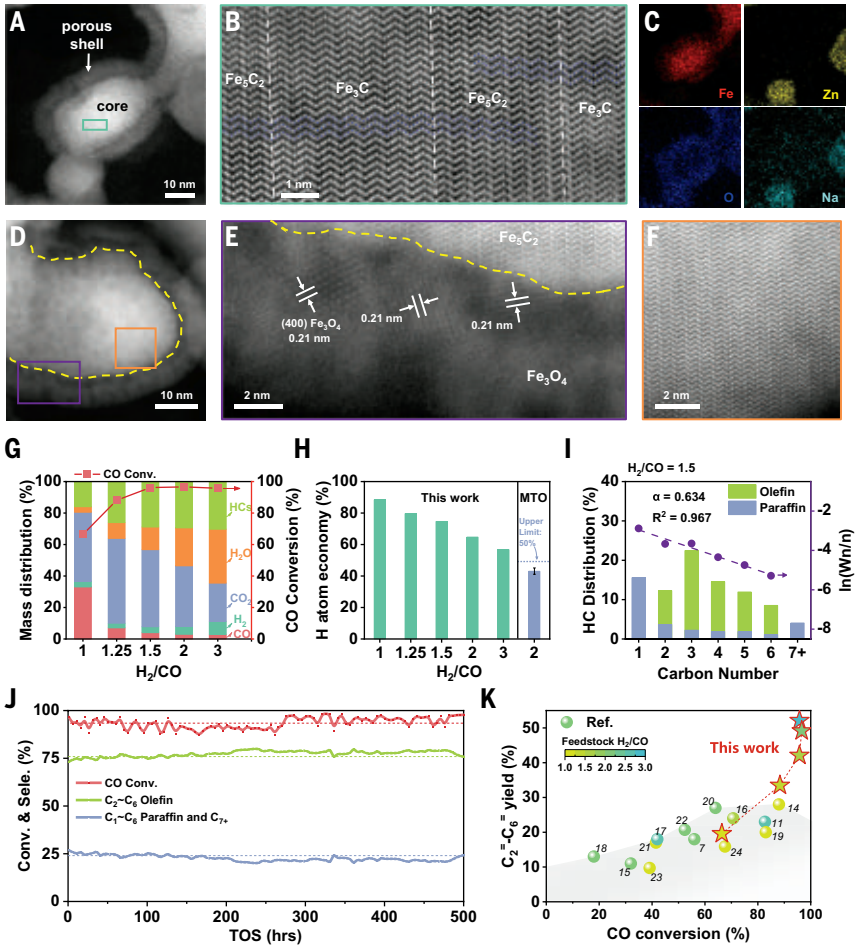


Fig. 1. Na/FeC_x@Fe_3O_4 catalyst with high HAE and olefin yield. (**A**) ADF-STEM image of a catalyst particle after 24 hours of reaction. (**B**) Magnified atomic ADF-STEM image of the core area marked by the blue frame in (A). (**C**) EDS elemental mapping of (A). (**D**) Atomic ADF-STEM image of another catalyst particle after 24 hours of reaction. (**E** and **F**) Magnified atomic ADF-STEM images of areas marked by the colored frames in (D). (**G**) Mass balance and CO conversion under different H_2/CO ratios. (**H**) HAE in MTO and under different H_2/CO ratios in STO. (**I**) Hydrocarbon (HC) product distribution and chain growth factor under feedstocks with H_2/CO ratios of 1.5 in 2 MPa at 623 K. (**J**) Hydrocarbon product distribution and CO conversion during the long-time test (623 K, 2 MPa, H_2/CO ratio of 2). TOS, time on stream. (**K**) Catalyst performance in this work (stars) versus previously reported metal-based Fischer-Tropsch catalysts (circles).

Transient analysis with mass spectrum also validated that the increased pressure could activate two processes simultaneously (Fig. 2C). In addition, the peak of CO_2 appeared earlier than the hydrocarbons, allowing the H_2 supply from WGS in situ for STO. Signals of CD_3H and CH_3D were observed in the mass spectra after isotope labeling by cofeeding of D_2O in syngas conversion (Fig. 2D). This result provided direct evidence that H from H_2O was incorporated into the hydrocarbon products (fig. S8).

We intentionally blocked the WGS path by introducing nonaqueous $MgCl_2$ powders. In situ characterization revealed that $MgCl_2$ adsorbed water at 623 K to form a MgOHCl-like phase (figs. S9 and S10) (*25*–*27*), directly cutting the water supply source in WGS reaction. The associated CO conversion markedly decreased from 96 to ~40% (Fig. 2E). The CO_2 selectivity dropped from 29.3 to 3.5% accordingly, validating the premise that CO_2 was mainly produced in the WGS reaction instead of in the STO reaction.

We used mass spectrometry to monitor the variation in H_2 and CO concentrations. The calculated H_2/CO ratio at the reactor outlet was ~7 versus the value of 2 in the fresh feedstock (fig. S11). The value of 7 would not be possible for the STO reaction alone because the formation of paraffins (C_nH_{2n+2}) and olefins (C_nH_{2n}) always consumes more H_2 than CO in the molar amount. Therefore, this much higher H_2/CO ratio had to come from the WGS-STO coupling effect.

In sharp contrast, blocking of the WGS path by adding 25 and 33 wt % $MgCl_2$ caused the H_2/CO ratios in the outlet to decrease to ~4.0 and ~2.9, respectively (Fig. 2E and figs. S11 and S12). We further confirmed that H_2 supply from WGS path contributed substantially to the STO reaction in producing olefins and paraffins (fig. S13). Long-term catalytic tests in which we added 25 wt % $MgCl_2$ indicated that both the CO conversion and CO_2 selectivity, as well as the H_2/CO ratio in the outlet, gradually recovered with the reaction time over the catalysts (Fig. 2F). Blocking the WGS pathway has a slightly promoting effect on the overall growth of the carbon chain and the O/P ratio in hydrocarbons (fig. S14). Similar phenomena were observed over the catalyst mixed with another typical desiccant, $CaCl_2$ (fig. S15) (*28*). In brief, blocking the WGS path with desiccants had only a temporary poisoning effect on the catalyst, which further confirmed the WGS-STO coupling effect. We concluded that the origin of the structure-performance relation was the catalyst core-shell structure that efficiently coupled WGS and STO (Fig. 2, G and H).

with that of previously reported metal-based Fischer-Tropsch catalysts (*7*, *11*, *14*–*24*). The present catalyst exhibited exceptional performance across a wide range of H_2/CO ratios (Fig. 1K). Specifically, within the H_2/CO ratio range of 1.25 to 2, the catalyst achieved a carbon-based yield of C_2 to C_6 olefins >30%, surpassing the previously reported values (table S4).

WGS-STO reaction coupling

We designed several experiments to study the hydrogen element flow contributing to high HAE. First, we showed that the as-prepared catalyst exhibited a high performance when fed with H_2O and CO at 623 K and 2 MPa (Fig. 2A and fig. S7). Under these conditions, H_2 can be produced only through the WGS reaction. Second, CO conversion was only 10% and HAE was only 23% at 523 K (Fig. 2B), but these increased substantially to 95% and 62% at 623 K, respectively. This result corroborated that high HAE was strongly temperature dependent, requiring concurrent activation of WGS and STO on the catalyst surface.

Coupling stoichiometric analysis for high HAE

Quantification of the contribution of WGS in syngas conversion is crucial for rationally designing catalysts that achieve high HAE. The simplified theoretical analysis was performed by assuming CO_2 was only produced in the WGS reaction (as confirmed by the experiments blocking the WGS reaction). In detail, the chemical equations for producing propylene as a single hydrocarbon were as follows.

STO reaction:

$$CO + 2H_2 = \frac{1}{3}C_3H_6 + H_2O \quad (1)$$

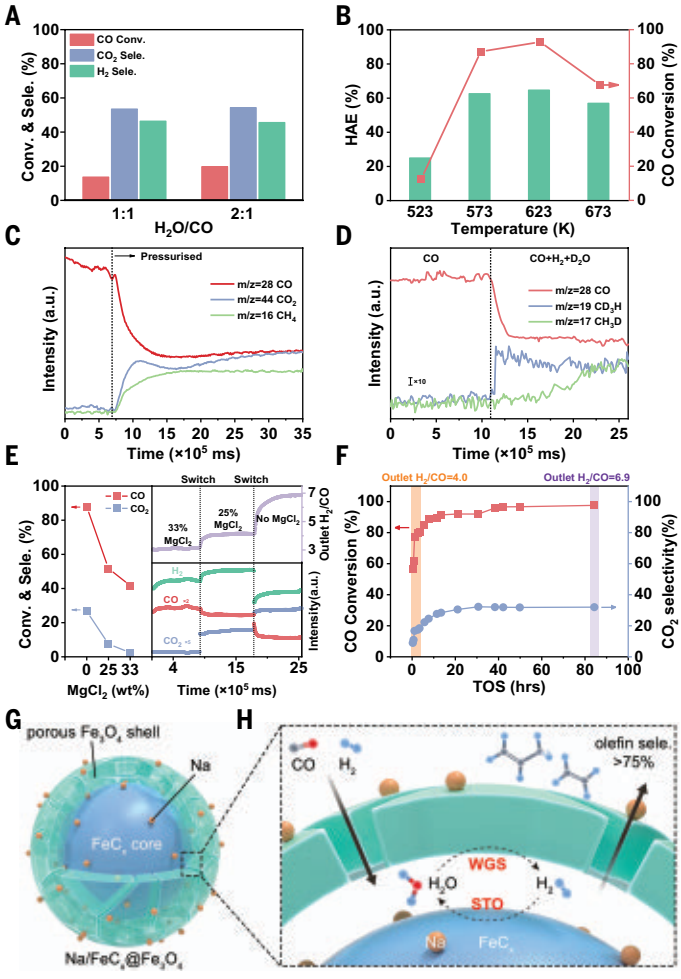


Fig. 2. Coupling of WGS and STO in the Na/FeC_x@Fe_3O_4 catalyst. (**A**) WGS capability at 623 K and 2 MPa. (**B**) CO conversion and HAE at various temperatures (2 MPa, feedstock with an H_2/CO ratio of 2). (**C**) Transient data obtained by a mass spectrometer for a pressurized reaction. Test conditions were 673 K and H_2/CO = 2. (**D**) Transient data obtained by mass spectrometry for isotopic tracing. Signals from CH_2D_2 (m/z = 18) or CD_4 (m/z = 20) were not captured because their signal intensity may be affected by H_2O or D_2O. (**E**) CO conversion and CO_2 selectivity obtained by blocking WGS using $MgCl_2$ absorbent at a feed H_2/CO ratio of 2, and outlet transient signal intensities obtained by multichannel micro-fixed-bed reactors coupling gas phase mass spectrometer switch experiments. (**F**) The long-term test and outlet H_2/CO ratios of the activated catalyst with the 25 wt % $MgCl_2$ adsorbent at a feed H_2/CO ratio of 2, where reaction initial (TOS = 2 hours) is shown in yellow and reaction restoration (TOS = 85 hours) in violet. (**G**) Schematic of the core-shell catalyst model comprising a Na/FeC_x syngas conversion catalyst. FeC_x particles are shown in blue and Na particles in brown. The Na/Fe_3O_4 WGS catalyst is shown in cyan. (**H**) Coupled WGS-STO reaction scheme.

WGS reaction:

$$xCO + xH_2O = xCO_2 + xH_2 \quad (2)$$

Thus, the overall reaction equation between Eq. 1 and Eq. 2 would be Eq. 3:

$$(1+x)CO + (2-x)H_2 = \frac{1}{3}C_3H_6 + (1-x)H_2O + xCO_2 \quad (3)$$

Here, the H_2/CO ratio needed was $(2 - x)/(1 + x)$, which was consistently lower than 2 because x should be a positive number. In this way, the equations to produce each type of single hydrocarbon (C_1 to C_4) were obtained (table S4).

Trends of the H_2/CO ratio needed for STO, HAE, and CO_2 yield were obtained by varying x (the contribution of WGS) (Fig. 3, A to D, and fig. S16). The higher the x value, the lower the H_2/CO ratio of the feedstock required and the higher the value of HAE achieved for a single hydrocarbon product. The theoretical curves for syngas conversion (producing a mixture of hydrocarbon species) were obtained with the feedstock of H_2/CO = 1.5 (Fig. 3E). The changing tendency of experimental plots was consistent with the theoretical ones (Fig. 3F and table S6). HAE values with practical product distribution were near those in the theoretical calculation.

We could calculate x values (the contribution of WGS) from the experimental data because the mole fraction of CO_2 (~24 to 43%) was equal to $x/(1 + x)$. The x value over the present catalyst was ~74% under an H_2/CO ratio of 1 but decreased to 34% under an H_2/CO ratio of 3. Rational design of the catalyst was possible by following this guideline. Considering most of the previous works intended to inhibit the WGS effect of the catalyst (*7*, *18*, *20*), our work opened a new door to use WGS positively and economically in the STO process. Furthermore, the fitting method proposed in this work was based on the stoichiometry of theoretical and actual reaction equations, exhibiting broad applicability to competitive and autocatalytic reaction systems.

Industrial process analysis

On the basis of the WGS-STO coupling effect, we built a coupled process flowchart for detailed analysis, benchmarking it against the already commercialized WGS-methanol-MTO flowchart (Fig. 4, A and B). The emission of CO_2 in the WGS and STO, steam consumed in the WGS, and wastewater produced in the STO were quantitatively evaluated based on a normalized unit capacity (tables S7 to S13). The higher HAE inherent to the WGS-STO flowchart enables olefin production at H_2/CO ratios <2. Although CO_2 emission was higher in the separate STO section the gross CO_2 emission was decreased in the integrated WGS-STO process under a low H_2/CO ratio in the feedstock (Figs. 1I and 4C). Concurrently, less steam was input in WGS section and less wastewater was produced in STO section for such an integrated WGS-STO process under a low H_2/CO ratio (Fig. 4D).

The environmental factor (E-factor) can be used to indicate atom consumption at the industrial scale (*29*, *30*). CO_2 and wastewater are the by-products. The E-factor of integrated WGS-STO increased with the H_2/CO ratio in feedstock, suggesting a higher atom consumption and lower economy (Fig. 4E). However, the complete E-factor (cE-factor) of WGS-STO route dropped by 46% compared with that of the WGS-methanol-MTO route because of the high HAE of the former case (*31*) (Fig. 4F and tables S14 to S16). Apparently, the present catalyst allowed the building of technology with water saving, low CO_2 emission, and high HAE by using the syngas feedstock with a low H_2/CO ratio to produce high-yield olefins, advancing process sustainability.

Discussion

We developed a catalyst with an in situ–formed Na/FeC_x@Fe_3O_4 core-shell NP structure, which coupled WGS and STO reactions to produce olefins with high HAE using syngas feedstock with low H_2/CO ratios. High CO conversion (95%), high olefin selectivity (>75%, hydrocarbon base), and high hydrocarbon yield (33 wt %, feedstock base) were obtained simultaneously by feeding a lower H_2/CO ratio feedstock at 623 K and 2 MPa. The HAE value with the coupled WGS-STO effect was between 66 and 86% for an H_2/CO ratio of 1 to 2, far higher than the values of 43 to 47% for the WGS-methanol-MTO route for an H_2/CO ratio of 2. The coupling effect of the catalyst depended on the amount of Na addition and temperature and remained stable after a continuous 500-hour test.

The WGS contribution was validated by adding desiccants ($MgCl_2$ or $CaCl_2$) to catalysts to block the WGS reaction path. We quantified the WGS contribution by fitting the experimental data with the

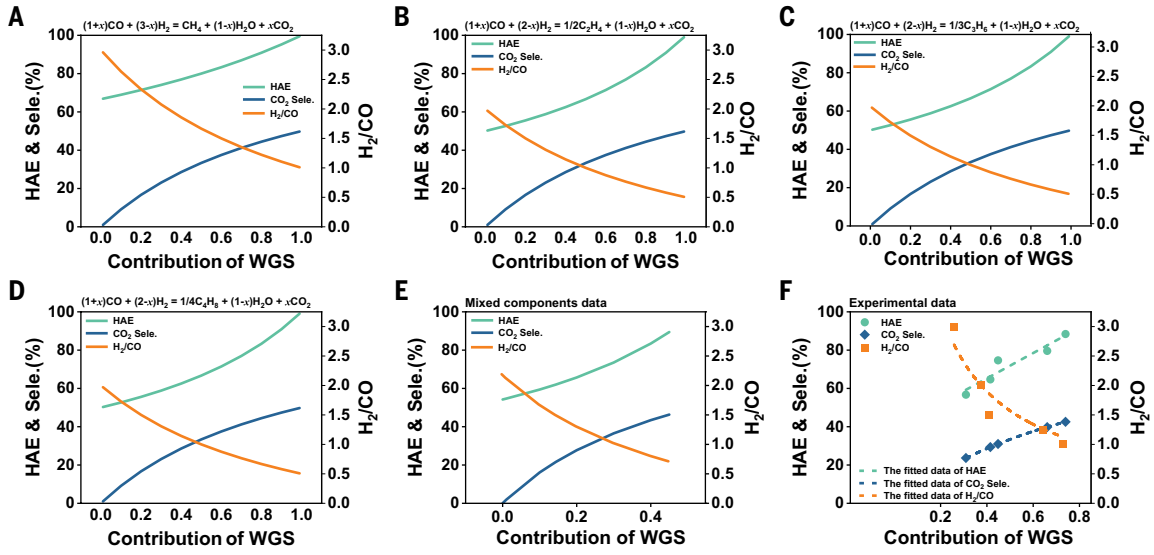


Fig. 3. Quantification of the contribution of WGS in syngas conversion. (**A** to **D**) Theoretical calculations of the variation of CO_2 selectivity, HAE, and H_2/CO ratio as a function of WGS proportion with different olefin and paraffin products. (**E** and **F**) Theoretical calculations of mixed-components data (E) and experimental data for syngas conversion (mixture of hydrocarbon products) (F).

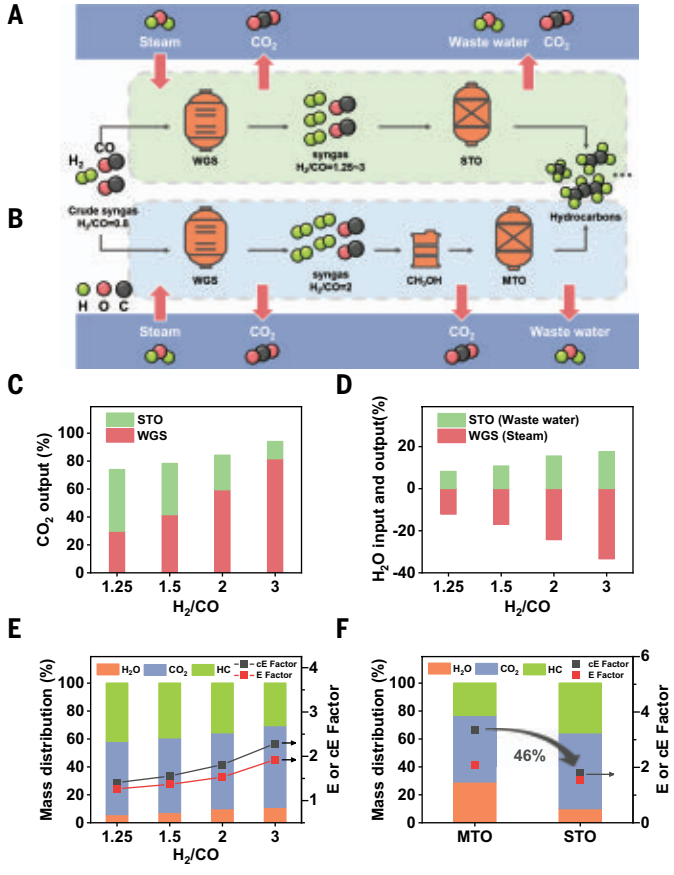


Fig. 4. Environmental factor analysis and comparison with MTO technology. (**A**) Flowsheet including WGS and STO (light green). (**B**) Flowsheet including WGS and MTO (light blue). (**C** and **D**) The normalized amount of CO_2 (C) and H_2O (D) stream under different H_2/CO ratios in the WGS-STO flowsheet. (**E**) Mass distribution and corresponding environmental factors in the series process of WGS-STO by varying the composition of the raw gas. (**F**) Comparison of MTO and STO processes with WGS on the mass distribution and environmental factors under the same feeding conditions shown in (A) and (B).

simulation trend. Industrial process analysis confirmed that the WGS-STO route with a lower H_2/CO ratio exhibited reduced steam consumption, decreased wastewater generation, lower CO_2 emissions, and higher HAE. Compared with the WGS-methanol-MTO route, the cE-factor of the WGS-STO route was reduced by 46%. The present catalyst with superior HAE for STO reaction provides a sustainable alternative olefins production route to the present MTO technology.

Materials and methods

Catalyst preparation

$Fe(NO_3)_3{\cdot}9H_2O$, $Zn(NO_3)_2{\cdot}6H_2O$, Na_2CO_3, $NH_3{\cdot}H_2O$, K_2CO_3, $MgCl_2$, and $CaCl_2$ were purchased from InnoChem Science & Technology Co., Ltd. and used as received. The $Na/FeC_x@Fe_3O_4$ catalyst was formed by carbonizing the precursor $NaZnFe_2O_4$. $NaZnFe_2O_4$ was synthesized by the coprecipitation method. In a typical experiment, 0.1 mol/L $Fe(NO_3)_3{\cdot}9H_2O$ and 0.1 mol/L $Zn(NO_3)_2{\cdot}6H_2O$ were thoroughly mixed and preheated above 333 K. Thereafter, a certain volume of 0.1 mol/L Na_2CO_3 was added dropwise to the mixture until the solution pH rose to ~10 and then aged for 12 hours. After straining and drying, the catalyst was sent to a muffle furnace after sufficient grinding. The temperature was raised from room temperature to 773 K at a rate of 2 K/min to get the $NaZnFe_2O_4$ catalyst sample. The $NH_3ZnFe_2O_4$ and $KZnFe_2O_4$ catalysts were obtained using the same method except that the sodium carbonate was replaced by ammonia and potassium carbonate, respectively.

Physically mixing with $MgCl_2$ powder in different mixing ratios was used for the activated catalyst (after 24 hours of reaction) that had completed the iron carbide phase transition under the nitrogen atmosphere. In this case, the influence of $MgCl_2$ on the iron carburization process could be excluded. A nitrogen atmosphere was used during the mixing process, and the two components were gradually stirred with a medicine spoon for 20 min to ensure that the mixture became homogeneous without the introduction of mechanical forces such as grinding. $CaCl_2$ was introduced using the same method.

The catalyst was dissolved in water for ultrasonic dispersion, underwent separate and repeat cleaning 10 times, and dried to obtain an $ZnFe_2O_4$ spinel catalyst without sodium. Sodium loading was confirmed by ICP-AES to be 0.08% decreased from ~45%.

Syngas to olefins reaction

Single-channel micro-fixed-bed reactors, multichannel micro-fixed-bed reactors, and 100-g fixed-bed reactors were used for catalyst evaluation. It had been proved by blank test that the reactor's alloy material has no catalysis effect on syngas feed (H_2/CO=3:1~1:1, with a certain amount of N_2). Before the reaction, the catalyst was reduced under the synthesis atmosphere for 2 hours at 0.1 MPa and 673 K.

All reaction products were analyzed by the Shimadzu GC-2014 gas chromatograph (GC). The GC includes an FID detector for detecting the contents of hydrocarbons separately and a TCD detector to measure the conversion and selectivity of by-products. Conversion was calculated based on the changes in CO contents in the inlet and outlet when N_2 was added as an internal standard. Product selectivity was calculated based on the product peak area and corresponding correction factors. Element balance calculations were performed to verify the accuracy of the results.

Screenshots of the results of random sampling after reaction with different feed gas ratios under chromatographic configuration are shown in fig. S17.

Calculation of mass balance, HAE, and E(cE)-factor

All distribution of carbon components and all hydrocarbon components, as well as the volume of H_2, were obtained from the data of GC. The water component was undetectable with GC. Carbon element balance (as CO, CO_2, and hydrocarbons) was first accurately determined. Then, the H element balance (as H_2 gas, hydrocarbon, and water) and O element balance (as CO, CO_2, H_2O) were determined. To validate the data from the micro-fixed-bed reactor, an independently designed filling max-value 100 g fixed-bed reactor was used, with the inner tubes of the bed using circulating water-water vapor for heat exchange, and the bed filled with 25 to 100 g of formed catalyst. The inlet flowmeter for inlet material calculation, liquid-phase condenser balance, and wet flowmeter for material calculation at the outlet can realize the macro material balance with an error of 5% or less.

The economy of hydrogen atoms demonstrates the sustainability of synthesis methods and processes, converting as much hydrogen as possible into the final target product during the reaction process. The formula for calculating HAE (%) is as follows:

$$\mathrm{HAE} = \frac{\text{Mass of H atoms in target product}}{\text{Mass of H atoms in reactant}} \times 100\% \quad (4)$$

The E-factor (which can be further abbreviated as EF) in Eq. 5 made contributions to achieving green chemistry, circular economy, and sustainable development based on resource efficiency and minimizing waste in design (*32*). It means a paradigm shift from focusing solely on process efficiency of chemical yield to optimizing resource efficiency and eliminating waste, with no specific units and is dry-based data.

$$\mathrm{EF} = \frac{\text{Total mass of waste, without } H_2O}{\text{Produced mass of the catalyst}} \times 100\% \quad (5)$$

Specifically, considering that any necessary water treatment (separation, purification, etc.) will have a negative impact on the competitiveness of the process, we have also calculated the cE-factor (which can be abbreviated as cEF) (*33*, *34*) in Eq. 6.

$$\mathrm{cEF} = \frac{\text{Total mass of waste, including } H_2O}{\text{Produced mass of the catalyst}} \times 100\% \quad (6)$$

The calculations in this work do not take into account the consumption or emissions due to the separation process but do illustrate the CO_2 emissions due to methanol refining and recycling during the MTO process.

Characterization

The atomic-scale STEM experiments were performed using a Cs-corrected scanning transmission electron microscope (FEI Titan Cubed Themis G2 300) operated at 300 kV with a convergence semi-angle of 23.6 mrad, and the beam current was <30 pA. The microscope was equipped with a DCOR+ spherical aberration corrector for the electron probe, which was aligned using a standard gold sample before the experiments. The dwell time of probe scanning was 32 μs for all the EDS mapping images. The atomic-resolution integrated differential phase contrast STEM was also operated at 300 kV with a convergence semiangle of 15 mrad. The beam current was set between 1 and 0.5 pA.

The D8 Advance x-ray diffractometer from Bruke (Germany) was used to analyze the phase and crystal structure of the catalyst. A Cu-Kα target was used, with an acceleration voltage of 40 kV and an acceleration current of 120 mA. The test was performed by continuous scanning. JADE 6.0 software was used for further analysis of the data. Because of the interference of catalyst signals, a pure $MgCl_2$ sample was used in the in situ XRD experiments. Water was introduced into the in situ chamber by passing N_2 through a bubbling device. To mimic the pretreatment of the mixture catalyst, the $MgCl_2$ sample was purged with dry N_2 at 673 K before the introduction of wet N_2.

The IRIS Intrepid II inductively coupled plasma atomic emission spectrometer (ICP-AES) from Thermo Corporation was used to analyze the content of elements in the catalyst body phase to obtain the ratio of metals and additives.

The instantaneous gas composition at the reactor outlet was monitored using a QGA 2.0 Gas Quantification Mass Spectrometer from Hiden Analytical Ltd. The unit was equipped with a high-performance external scroll pump to counteract the effects of high H_2 levels in the system gas. Data were analyzed using MAsoft Pro software. The mass spectrometry conversion monitoring experiments for $MgCl_2$-doped catalysts were conducted using a multichannel fixed-bed reactor, which is equipped with a solenoid valve at its outlet for flow path switching.

REFERENCES AND NOTES

1. H. M. Torres Galvis, K. P. de Jong, Catalysts for production of lower olefins from synthesis gas: A review. *ACS Catal.* **3**, 2130–2149 (2013). doi: 10.1021/cs4003436
2. P. Zhai *et al.*, Development of direct conversion of syngas to unsaturated hydrocarbons based on Fischer-Tropsch route. *Chem* **7**, 3027–3051 (2021). doi: 10.1016/j.chempr.2021.08.019
3. F. Jiao *et al.*, Selective conversion of syngas to light olefins. *Science* **351**, 1065–1068 (2016). doi: 10.1126/science.aaf1835; pmid: 26941314
4. F. Jiao *et al.*, Disentangling the activity-selectivity trade-off in catalytic conversion of syngas to light olefins. *Science* **380**, 727–730 (2023). doi: 10.1126/science.adg2491; pmid: 37200424
5. Y. Lu, T. Lee, Influence of the feed gas composition on the Fischer-Tropsch synthesis in commercial operations. *J. Nat. Gas Chem.* **16**, 329–341 (2007). doi: 10.1016/S1003-9953(08)60001-8
6. S. Yang *et al.*, An integrated framework for modeling, synthesis, analysis, and optimization of coal gasification-based energy and chemical processes. *Ind. Eng. Chem. Res.* **51**, 15763–15777 (2012). doi: 10.1021/ie3015654
7. Y. Xu *et al.*, A hydrophobic FeMn@Si catalyst increases olefins from syngas by suppressing C1 by-products. *Science* **371**, 610–613 (2021). doi: 10.1126/science.abb3649; pmid: 33542132
8. P. Tian, Y. Wei, M. Ye, Z. Liu, Methanol to olefins (MTO): From fundamentals to commercialization. *ACS Catal.* **5**, 1922–1938 (2015). doi: 10.1021/acscatal.5b00007
9. G. Glenk, S. Reichelstein, Economics of converting renewable power to hydrogen. *Nat. Energy* **4**, 216–222 (2019). doi: 10.1038/s41560-019-0326-1
10. J. Zhu *et al.*, Dynamic structural evolution of iron catalysts involving competitive oxidation and carburization during CO_2 hydrogenation. *Sci. Adv.* **8**, eabm3629 (2022). doi: 10.1126/sciadv.abm3629; pmid: 35119927
11. P. Zhai *et al.*, Highly tunable selectivity for syngas-derived alkenes over zinc and sodium-modulated Fe_5C_2 catalyst. *Angew. Chem. Int. Ed.* **55**, 9902–9907 (2016). doi: 10.1002/anie.201603556; pmid: 27445106
12. I. Puskas, R. S. Hurlbut, Comments about the causes of deviations from the Anderson–Schulz–Flory distribution of the Fischer–Tropsch reaction products. *Catal. Today* **84**, 99–109 (2003). doi: 10.1016/S0920-5861(03)00305-5

13. M. E. Dry, The Fischer–Tropsch process: 1950–2000. *Catal. Today* **71**, 227–241 (2002). doi: 10.1016/S0920-5861(01)00453-9
14. H. M. Torres Galvis *et al.*, Supported iron nanoparticles as catalysts for sustainable production of lower olefins. *Science* **335**, 835–838 (2012). doi: 10.1126/science.1215614; pmid: 22344440
15. L. Zhong *et al.*, Cobalt carbide nanoprisms for direct production of lower olefins from syngas. *Nature* **538**, 84–87 (2016). doi: 10.1038/nature19786; pmid: 27708303
16. P. Wang *et al.*, Efficient conversion of syngas to linear α-olefins by phase-pure χ-Fe_5C_2. *Nature* **635**, 102–107 (2024). doi: 10.1038/s41586-024-08078-5; pmid: 39415021
17. F. Qian *et al.*, Stabilized ε-Fe_2C catalyst with Mn tuning to suppress C1 byproduct selectivity for high-temperature olefin synthesis. *Nat. Commun.* **15**, 5128 (2024). doi: 10.1038/s41467-024-49472-x; pmid: 38879628
18. J. Xie *et al.*, Promoted cobalt metal catalysts suitable for the production of lower olefins from natural gas. *Nat. Commun.* **10**, 167 (2019). doi: 10.1038/s41467-018-08019-7; pmid: 30635560
19. C. Wang *et al.*, Fischer-Tropsch synthesis to olefins boosted by MFI zeolite nanosheets. *Nat. Nanotechnol.* **17**, 714–720 (2022). doi: 10.1038/s41565-022-01154-9; pmid: 35817859
20. W. Fang *et al.*, Physical mixing of a catalyst and a hydrophobic polymer promotes CO hydrogenation through dehydration. *Science* **377**, 406–410 (2022). doi: 10.1126/science.abo0356; pmid: 35862543
21. Y. Liu, J.-F. Chen, J. Bao, Y. Zhang, Manganese-modified Fe_3O_4 microsphere catalyst with effective active phase of forming light olefins from syngas. *ACS Catal.* **5**, 3905–3909 (2015). doi: 10.1021/acscatal.5b00492
22. Y. Xu *et al.*, Effects of surface hydrophobization on the phase evolution behavior of iron-based catalyst during Fischer-Tropsch synthesis. *Nat. Commun.* **15**, 7099 (2024). doi: 10.1038/s41467-024-51472-w; pmid: 39154082
23. O. Zhuo *et al.*, Stabilizing the active phase of iron-based Fischer-Tropsch catalysts for lower olefins: Mechanism and strategy. *Chem. Sci.* **10**, 6083–6090 (2019). doi: 10.1039/C9SC01210A; pmid: 31360413
24. B. Gu *et al.*, Effects of the promotion with bismuth and lead on direct synthesis of light olefins from syngas over carbon nanotube supported iron catalysts. *Appl. Catal. B* **234**, 153–166 (2018). doi: 10.1016/j.apcatb.2018.04.025
25. S. Kashani-Nejad, K. W. Ng, R. Harris, Preparation of MgOHCl by controlled dehydration of $MgCl_2{\cdot}6H_2O$. *Metall. Mater. Trans., B, Process Metall. Mater. Proc. Sci.* **35**, 405–406 (2004). doi: 10.1007/s11663-004-0042-4
26. G. Balakrishnan, R. Velavan, K. Mujasam Batoo, E. H. Raslan, Microstructure, optical and photocatalytic properties of MgO nanoparticles. *Results Phys.* **16**, 103013 (2020). doi: 10.1016/j.rinp.2020.103013
27. N. Klammer, C. Engtrakul, Y. Zhao, Y. Wu, J. Vidal, Method to determine MgO and MgOHCl in chloride molten salts. *Anal. Chem.* **92**, 3598–3604 (2020). doi: 10.1021/acs.analchem.9b04301; pmid: 32073829
28. G. C. Sinke, E. H. Mossner, J. L. Curnutt, Enthalpies of solution and solubilities of calcium chloride and its lower hydrates. *J. Chem. Thermodyn.* **17**, 893–899 (1985). doi: 10.1016/0021-9614(85)90083-7
29. R. A. Sheldon, Metrics of green chemistry and sustainability: Past, present, and future. *ACS Sustain. Chem.& Eng.* **6**, 32–48 (2017). doi: 10.1021/acssuschemeng.7b03505
30. R. A. Sheldon, The E factor 25 years on: The rise of green chemistry and sustainability. *Green Chem.* **19**, 18–43 (2017). doi: 10.1039/C6GC02157C
31. F. Roschangar, R. A. Sheldon, C. H. Senanayake, Overcoming barriers to green chemistry in the pharmaceutical industry – the Green Aspiration Level™ concept. *Green Chem.* **17**, 752–768 (2015). doi: 10.1039/C4GC01563K
32. R. A. Sheldon, The E factor at 30: A passion for pollution prevention. *Green Chem.* **25**, 1704–1728 (2023). doi: 10.1039/D2GC04747K
33. D. P. Debecker *et al.*, Shaping effective practices for incorporating sustainability assessment in manuscripts submitted to *ACS Sustainable Chemistry & Engineering*: Catalysis and catalytic processes. *ACS Sustain. Chem.& Eng.* **9**, 4936–4940 (2021). doi: 10.1021/acssuschemeng.1c02070
34. P.-I. Dassie *et al.*, Coupling of solvent-free synthesis and reactive extrusion of alumina: An ecologically efficient integration for heterogenous catalyst synthesis. *Green Chem.* **25**, 2800–2814 (2023). doi: 10.1039/D2GC04714D

ACKNOWLEDGMENTS

We thank G. Liu (Beijing University of Technology) and Z. Li, X. Sui, and P. Li (Ordos Laboratory) for their helpful contributions to catalyst evaluation and characterization. **Funding:** This work was supported by the National Key Research and Development Program of China (2024YFB4105403) and the National Key Research and Development Program of China (2024YFB4105402). **Author contributions:** W.Q. initiated and led the whole project. W.Q. and C.C. supervised and coordinated the project. C.G. and W.S. contributed to catalyst development and characterization. H.W. and X.C. performed the STEM characterization. W.H., N.Y., Y.Y., S.Y., H.L., and M.M. participated in catalyst evaluation. X.L. and R.Z. assisted in catalyst preparation. C.G., C.C., and W.Q. wrote the manuscript. All authors discussed the results, edited the manuscript, and approved the final version. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.aea0774
Figs. S1 to S17; Tables S1 to S16; References (*35*, *36*)

Submitted 29 June 2025; accepted 8 September 2025

10.1126/science.aea0774

MOLECULAR BIOLOGY

Human RPA is an essential telomerase processivity factor for maintaining telomeres

Sourav Agrawal[1]†, Xiuhua Lin[1]†, Vivek Susvirkar[1]†, Michael S. O'Connor[2,3], Bianca L. Chavez[1], Victoria R. Tholkes[1], Grace P. Tauber[1], Qixiang He[1], Kaitlyn M. Abe[1], Xuhui Huang[2,3], Ci Ji Lim[1]*

Telomerase counteracts telomere shortening by repeatedly adding DNA repeats to chromosome ends. We identified the replication protein A (RPA) heterotrimer as a telomerase processivity factor critical for telomere maintenance. RPA stimulates telomerase processivity in vitro, and AlphaFold modeling predicts that RPA engages a telomerase surface distinct from the one bound by the shelterin subunit TPP1. Guided by these predictions, we engineered separation-of-function telomerase reverse transcriptase (TERT) mutants and found that the loss of RPA-mediated stimulation impairs telomere elongation, even when TPP1–POT1–mediated stimulation remains intact. Furthermore, short-telomere disease–associated TERT mutations reduce RPA-dependent telomerase stimulation, revealing a mechanistic link between impaired processivity and telomeropathies. Together, our findings establish human RPA as a key regulator of telomerase and offer molecular insights into telomere-related disease mechanisms.

Telomeres, the natural ends of linear chromosomes, are crucial for maintaining genome stability (*1*). In humans, telomere maintenance (*2*, *3*) involves the extension of the G-rich single-stranded DNA (ssDNA) by telomerase (*4*–*7*) and the synthesis of the complementary C-strand by DNA polymerase α-primase (Polα-primase) in conjunction with the CTC1-STN1-TEN1 (CST) complex (*8*, *9*). Telomerase is recruited to telomeres by the shelterin protein TPP1 (*10*–*12*), which together with POT1 also stimulates telomerase repeat addition processivity (RAP) in vitro (*13*). However, the role of this stimulatory function in telomere maintenance in vivo remains unclear.

Replication protein A (RPA) is a highly conserved heterotrimeric complex involved in DNA replication and repair (*14*). In ciliates and yeast, RPA-like complexes are implicated in telomerase regulation and telomere maintenance (*15*–*19*). In human cell extracts, RPA has been reported to modulate telomerase activity and processivity (*20*). However, whether RPA directly stimulates telomerase processivity and supports telomere maintenance in human cells remained unknown.

RPA stimulates telomerase processivity through ssDNA and TERT interactions

Using primer-extension assays, we found that recombinant human RPA (fig. S1) increased telomerase RAP approximately sixfold, an effect similar to although less potent than that of the TIN2-TPP1-POT1 (TPT) complex (Fig. 1, A and B). Like TPT (*13*), RPA specifically stimulated RAP without affecting overall telomerase activity (Fig. 1C). This function required the complete heterotrimer because individual subunits and subcomplexes were inactive (Fig. 1D and fig. S2). These results establish the RPA heterotrimer as a telomerase processivity factor in vitro.

To understand the mechanism of RPA stimulation, we investigated the role of RPA's ssDNA binding domains (Fig. 1D) (*14*). Deletion of the OB-A and OB-B domains, which are crucial for ssDNA binding (*21*), abolished the stimulatory effect, whereas removal of the OB-F domain involved in protein-protein interactions (*14*) did not (Fig. 1E and fig. S2). This suggests that RPA's interaction with ssDNA is critical for stimulating telomerase RAP. To further test the importance of ssDNA binding, we used a polyT oligonucleotide (T36) as a competitor. Titration of T36 sequestered RPA and led to a dose-dependent reduction in telomerase RAP stimulation by RPA but not by TPT (Fig. 1F). This differential effect occurs because RPA binds nonspecifically to any ssDNA (*14*), including the polyT competitor, whereas TPT (POT1) requires a specific telomeric sequence for binding through its POT1 subunit (*22*, *23*). This result further highlights the necessity of RPA's ssDNA binding for telomerase RAP stimulation.

TPP1–POT1 requires both protein-protein and protein-DNA interactions for telomerase RAP stimulation (*24*). We therefore examined whether RPA also has direct interactions with telomerase reverse transcriptase (TERT). We used AlphaFold2 multimer (*25*) to screen for potential protein-protein interactions between RPA subunits and TERT. AlphaFold2 predicted that the RPA32 OB-D domain interacts with TERT TEN domain (Fig. 1, G to I).

To test this prediction, we performed stoichiometric titrations of RPA to a fixed 1 μM DNA primer concentration in the direct telomerase assay. By titrating RPA from substoichiometric to equimolar levels, we assessed whether telomerase processivity enhancement by RPA arises purely from ssDNA binding or involves an additional TERT–RPA interaction, as predicted with AlphaFold2 and demonstrated for TPP1–POT1 (*24*). A purely ssDNA binding–dependent mechanism would yield a linear increase in processivity with increasing RPA concentration, whereas an additional TERT–RPA interaction should produce a nonlinear response.

We observed a nonlinear, hyperbolic increase in telomerase RAP, with near-maximal stimulation at substoichiometric RPA binding concentrations (Fig. 1, J and K). This response resembles the in vitro stimulation reported for TPP1–POT1 (*24*) and is consistent with the AlphaFold2 prediction that RPA directly interacts with TERT. A separate binding assay by using fluorescence polarization under the same conditions showed a linear correlation between RPA concentration and DNA primer occupancy (Fig. 1K), indicating that the nonlinear processivity enhancement is not due to binding inefficiency or miscalculated protein concentration.

To test for an in vivo interaction, we performed coimmunoprecipitation (co-IP) assays using mCherry-tagged baits (Fig. 1L and fig. S3). Tagged RPA subunits coprecipitated TERT, unlike the mCherry-RAP1 negative control. Specifically, mCherry-RPA14 produced a robust signal comparable with that of the TPP1 positive control, whereas the pull-down by mCherry-RPA32 was significant but more modest, a difference we attribute to potential steric hindrance from its N-terminal tag. Furthermore, co-overexpression of all three subunits with the tag on RPA14 [superRPA (*26*)] reduced the pull-down signal, likely because of competition from the elevated level of RPA complexes. RPA70 could not be assessed owing to poor expression. This pull-down by multiple RPA baits, which is consistent with mass spectrometry data detecting all three RPA subunits in purified telomerase holoenzyme (*27*), provides strong evidence for an in vivo interaction between telomerase and the RPA heterotrimer.

Together, these results demonstrate that RPA stimulates telomerase RAP through both ssDNA binding and protein-protein interactions, similar to TPP1–POT1.

[1]Department of Biochemistry, University of Wisconsin–Madison, Madison, WI , USA. [2]Department of Chemistry, University of Wisconsin–Madison, Madison, WI , USA. [3]Theoretical Chemistry Institute, University of Wisconsin–Madison, Madison, WI , USA. ***Corresponding author. Email: ciji.lim@wisc.edu** †These authors contributed equally to this work.

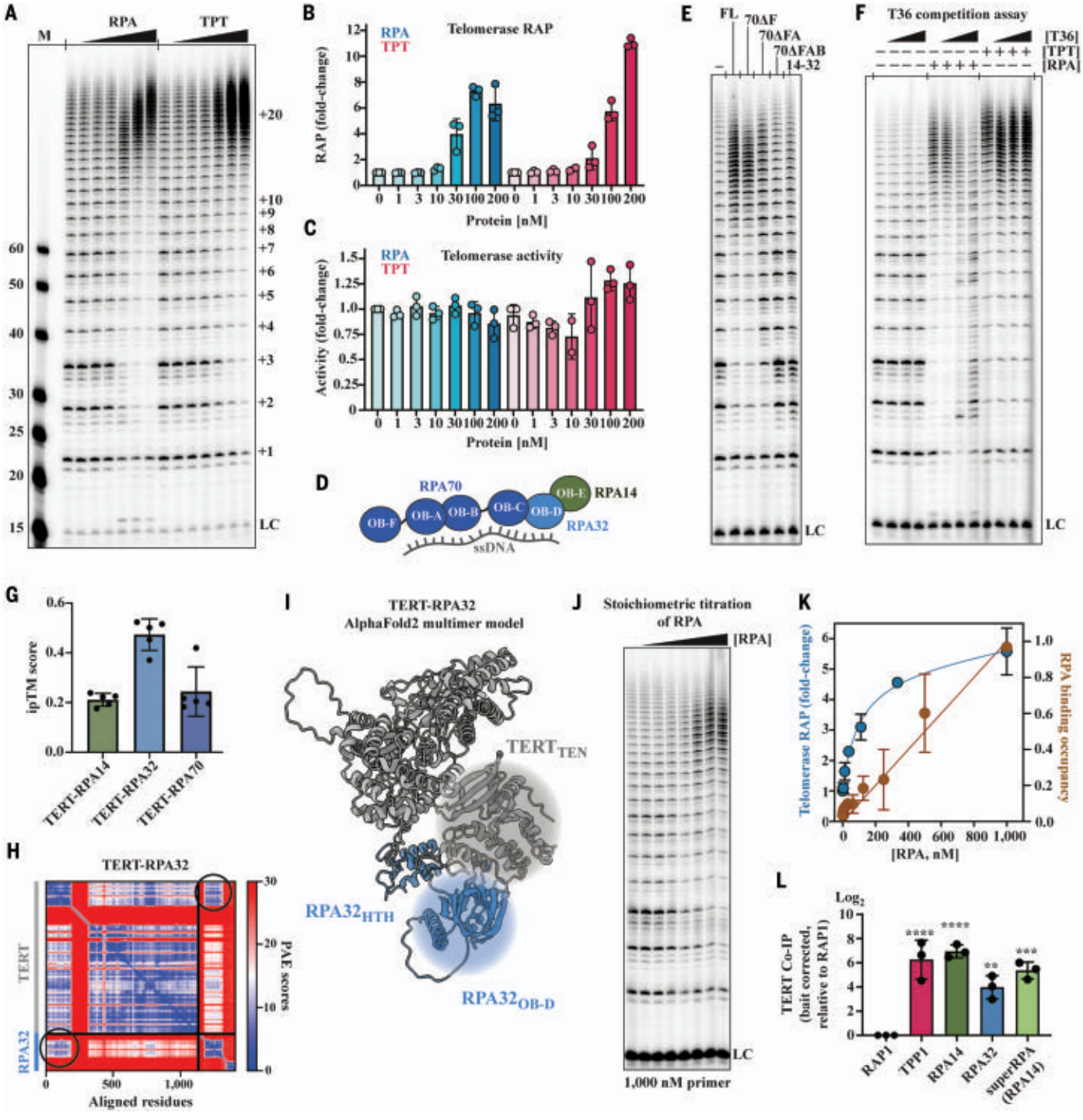


Fig. 1. Human RPA stimulates telomerase processivity through ssDNA and TERT interactions. (**A**) Direct telomerase assay showing repeat addition products from reactions with increasing concentrations of RPA or TPT (TIN2-TPP1-POT1). Numbers at right indicate the number of telomere repeats added. M, molecular weight markers; LC, loading control. (**B**) Quantification of telomerase RAP from experiments as in (A), shown as fold-change relative to the no-protein control. (**C**) Quantification of total telomerase activity from experiments as in (A), shown as fold-change. (**D**) Schematic of the human RPA heterotrimer highlighting its known ssDNA binding OB-fold domains (*14*). (**E**) Direct telomerase assay comparing full-length (FL) RPA with various RPA70 truncation mutants and the RPA14-32 heterodimer. Protein and protein complexes are used at 200 nM. (**F**) Competition assay showing the effect of a nontelomeric dT_{36} oligo (T36) on telomerase stimulation by TPT and RPA. (**G**) AlphaFold2-predicted integrated TM-score (ipTM) for TERT interactions with individual RPA subunits. (**H**) Predicted aligned error (PAE) plot for the TERT–RPA32 interaction, showing the expected distance error (angstroms) between residue pairs. Lower values (blue) indicate higher confidence. (**I**) AlphaFold2 multimer model showing the predicted interaction between the TERT TEN domain (gray) and the RPA32 OB-D domain (blue). (**J**) Stoichiometric titration of RPA (0 to 1000 nM protein concentration) in a direct telomerase assay using 1000 nM primer. (**K**) Quantification of telomerase RAP from (J) (blue, left *y* axis) and RPA binding occupancy determined by means of fluorescence polarization (brown, right *y* axis). (**L**) Co-IP results quantifying TERT pull-down by mCherry-tagged baits, normalized to bait input (mCherry–bait/mCherry–RAP1) and expressed relative to the RAP1 negative control. In all graphs, error bars indicate SD from $n = 3$ independent experiments. $^{****}P \leq 0.0001$; $^{***}P \leq 0.001$; $^{**}P \leq 0.01$.

RPA32 OB-D and TPP1 OB domains bind telomerase at separate sites

To determine whether RPA and TPP1 engage the same surface on telomerase, we docked our AlphaFold2-predicted $TERT_{TEN}$–RPA32 model onto a cryo–electron microscopy (cryo-EM) structure of TPP1-bound telomerase (Fig. 2A). This analysis revealed two distinct, nonoverlapping interfaces: RPA32 contacts the TEN domain, whereas the TPP1 OB domain spans both the TEN and IFD–TRAP regions, as previously described (*28*–*30*). The prediction of this distinct RPA32 interface remained consistent across models of increasing complexity, from AlphaFold2 simulations including TPP1 OB domain (fig. S4) to AlphaFold3 (*31*) predictions of the entire telomerase-RPA-TPP1-POT1-ssDNA assembly (fig. S5).

The predicted separation of the RPA32 and TPP1 binding sites suggested that we could create separation-of-function (SOF) TERT mutants to probe the distinctive functional roles of each factor (Fig. 2B). To test the stability of the predicted interactions, we performed all-atom molecular dynamics simulations. The simulations showed that highly conserved TERT residues (E79, R83, R87, T116, T117, D146, and D147) maintained stable contacts with RPA32 (Fig. 2, C to E, and fig. S6A), raising our confidence in selecting them for mutagenesis. (Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. In the mutants, other amino acids were substituted at certain locations; for example, E79R indicates that glutamate at position 79 was replaced by arginine.)

We reconstituted telomerase [by overexpressing TERT and telomerase RNA (TR) in human embryonic kidney (HEK) 293T cells] with a panel of TERT mutations at the predicted RPA-binding interface (E79A, R83A, R87A, T116A/T117A, and D146A/D147A), hereafter termed $TERT_{RPA}$. In direct telomerase assays, all $TERT_{RPA}$ mutants retained wild-type (WT) basal activity. However, three mutants—R83A, T116A/T117A, and D146A/D147A—selectively abolished RPA-mediated processivity stimulation while retaining stimulation by TPT (Fig. 2F and fig. S7), defining them as SOF mutants for RPA stimulation.

As a complementary approach, we tested previously reported $TERT_{TPT}$ mutants, K78E (*32*) and Y772A (*28*), which disrupt the TERT–TPP1 interaction. As predicted with AlphaFold modeling, both $TERT_{TPT}$ mutants disrupted TPT-mediated processivity stimulation, whereas RPA-mediated stimulation was only weakly affected (Fig. 2F). Y772A had little to no effect, whereas K78E caused a modest but detectable reduction in RPA stimulation, possibly because of its proximity to the predicted RPA32–$TERT_{TEN}$ interface. These findings further support the model that RPA32 and TPP1 OB domains bind distinct surfaces on TERT.

We next sought to validate the interaction by mutating the predicted contact residues on RPA32. Guided by the compact (~940 $Å^2$) interface, we introduced alanine substitutions at six predicted and conserved contact residues in the RPA32 OB-D domain (T88, W107, H131, R133, F135, and F144) (fig. S6B), which pair in the model with TERT residues D147, R83, T116/T117, E79, Q73, and T117, respectively (Fig. 2, C to E). The model predicted a key π-cation interaction between TERT R83 and RPA32 W107. Direct telomerase assays revealed a spectrum of effects on processivity (Fig. 2G and fig. S8); the W107A mutation markedly reduced RPA stimulation, whereas T88A, F135A, and F144A caused more modest losses, and H131A and R133A had little to no negative impact.

To ensure these effects were specific to the TERT–RPA32 interaction, we confirmed (by means of mass photometry) that each mutant RPA complex purified as an intact heterotrimer and (by means of fluorescence polarization assay) bound telomeric ssDNA with near-WT affinity (fig. S9). These results suggest that the observed loss in RPA stimulation is consistent with disruptions at the predicted binding interface, rather than protein misfolding or impaired DNA binding. Together, our findings indicate that several residues on the RPA32 OB-D surface contribute to telomerase stimulation.

RPA localizes to telomeres by means of telomerase without triggering a DNA damage response

We next used immunofluorescence (IF) to validate the TERT–RPA32 interaction in cells. We found that endogenous RPA32 colocalized with telomeres upon transient expression of both TERT and TR, indicating that telomerase is required to recruit or stabilize RPA at telomeres (Fig. 3A). To test whether the predicted interface was necessary, we analyzed cells expressing $TERT_{RPA}$ mutants (E79A, R83A, R87A, T116A/T117A, and D146A/D147A). All tested mutants reduced RPA signal at telomeres, with single-residue substitutions lowering colocalization by ~50% and double-residue substitutions abolishing it almost completely (Fig. 3B), demonstrating the TERT–RPA32 interaction is important for RPA accumulation at telomeres.

Because the $TERT_{RPA}$ mutations are distant from the known TERT–TPP1 interaction site (*28*, *29*), we next sought to confirm that they do not affect telomerase recruitment. Our IF imaging showed that key $TERT_{RPA}$ mutants, including R83A and T116A/T117A, were indeed recruited to telomeres as efficiently as WT TERT (Fig. 3C and fig. S10). Because these mutants retain intact recruitment but specifically lack RPA-mediated stimulation, they are true SOF mutants.

By contrast, the D146A/D147A mutation was an exception. Despite its distance from the TERT–TPP1 interaction site (fig. S11, A and B), the mutant exhibited a telomere recruitment defect, causing telomerase to accumulate in Cajal bodies (fig. S11C). However, it retained normal TPP1–POT1–mediated stimulation in vitro (Fig. 2F and fig. S7). How this mutant uncouples these two TPP1-mediated functions presents an interesting mechanistic question that warrants future investigation.

To confirm the TERT–RPA32 interaction at physiological levels, we performed a proximity ligation assay (PLA) (*33*) using probes for endogenous TR and RPA32. The assay detected discrete nuclear foci in HeLa cells, indicating that endogenous telomerase and RPA are in close proximity (Fig. 3, D and E). The PLA signal increased approximately threefold in cells expressing TERT from the AAVS1 safe-harbor locus (*34*), which is consistent with the estimated upper limit of fully assembled telomerase RNPs (approximately fourfold above parental) (*35*). Introducing the T116A/T117A mutation reduced the PLA signal by ~50 to 60%, confirming that the TERT–RPA32 interface mediates this proximity in cells.

Our biochemical assays established a dual requirement for RPA-mediated stimulation: TERT interaction and ssDNA primer engagement. To determine whether this dual requirement applies to RPA localization in cells, we analyzed a catalytically dead TERT mutant (D712A) (*36*) that prevents the formation of a nascent 3′ overhang for RPA to engage (fig. S12A). IF analysis showed that this mutant abolished RPA32 accumulation at telomeres, an outcome that mirrors the effect of our interface-disrupting mutants (fig. S12, B and C). This demonstrates that stable RPA localization at telomeres requires engagement with both TERT and telomeric ssDNA, confirming the dual requirement in a cellular context.

We investigated whether telomerase-mediated RPA localization triggers a DNA damage response (DDR), which occurs when RPA accumulates at telomeres upon POT1 depletion (*37*). We checked for two DDR markers: RPA32 phosphorylation (*38*) and 53BP1 foci (*37*). Unlike etoposide-treated controls, telomerase-overexpressing cells showed no RPA32 phosphorylation band shift on Western blots (Fig. 3F). Similarly, telomerase-overexpressing cells showed no increase in 53BP1 foci at telomeres (Fig. 3G and fig. S13). Together, these results demonstrate that telomerase-mediated RPA recruitment to telomeres occurs without activating DDR.

TERT SOF RPA mutants are defective in telomere maintenance in cells

To investigate RPA's physiological function in telomere maintenance, we used our SOF $TERT_{RPA}$ mutants (R83A and T116A/T117A), which retain telomerase recruitment to telomeres (Fig. 3C) but lack RPA-mediated

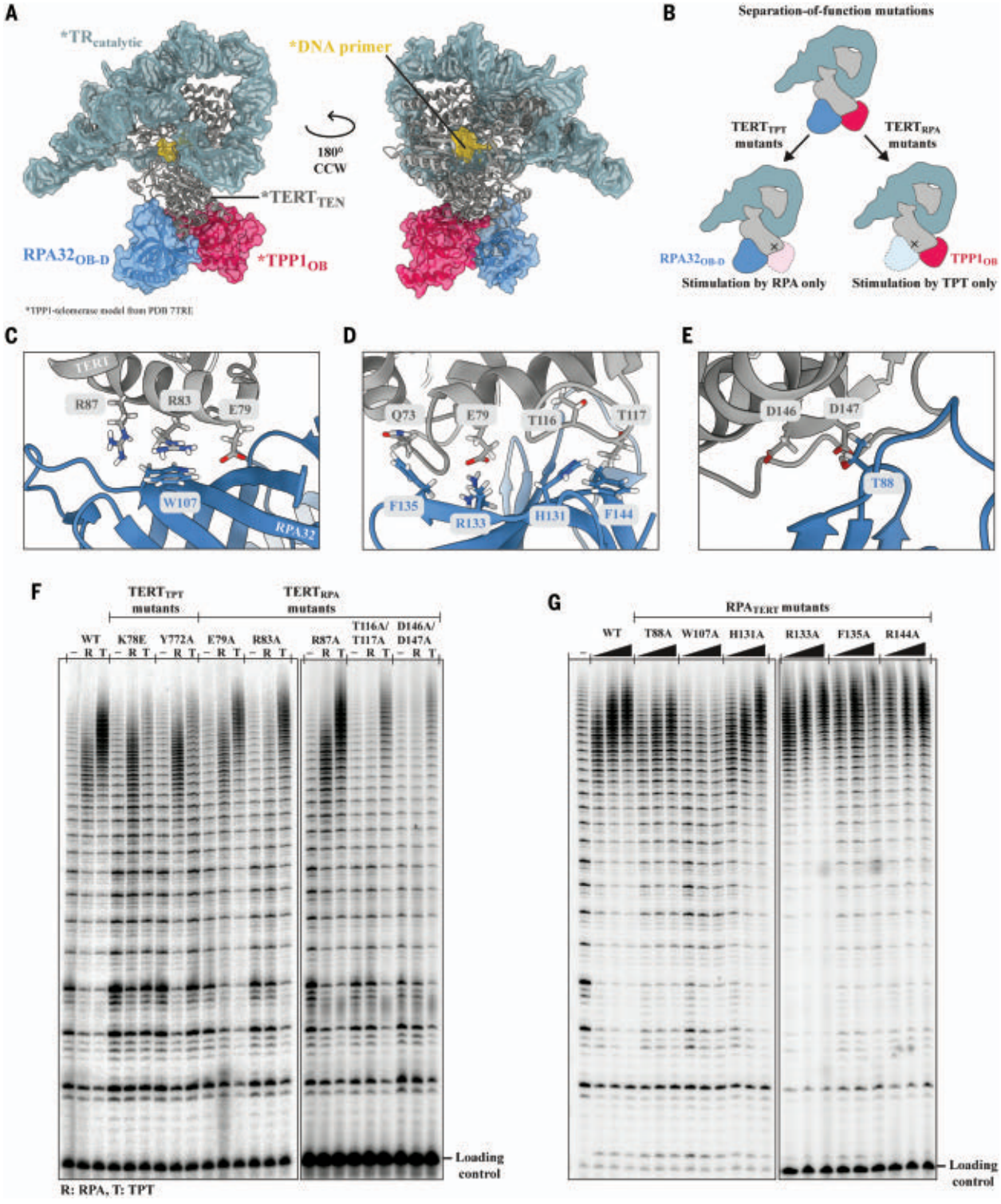


Fig. 2. Distinct TERT binding sites for RPA32 OB-D and TPP1 OB domains enable SOF mutations. (**A**) Hybrid structural model, shown in two views rotated 180° counterclockwise (CCW), showing the AlphaFold-predicted RPA32 OB-D domain (blue) docked onto the TPP1-bound telomerase cryo-EM structure [Protein Data Bank (PDB) ID 7TRE; asterisked components]. This illustrates the non-overlapping binding sites of RPA32 (blue) and TPP1 (red) OB domains on the TERT TEN domain. (**B**) Schematic illustrating the principle of TERT SOF mutants that are selectively stimulated by either RPA or TPT. (**C** to **E**) Close-up views of the predicted TERT–RPA32 interface highlighting key interacting residues. (**F**) Direct telomerase assay comparing WT TERT with $TERT_{TPT}$ mutants (K78E and Y772A) and $TERT_{RPA}$ mutants (E79A, R83A, R87A, T116A/T117A, D146A/D147A). Reactions were performed with no added protein (–), 200 nM RPA (R), or 200 nM TPT (T). (**G**) Direct telomerase assay testing WT TERT activity with increasing concentrations (50, 100, and 200 nM) of RPA complexes containing the indicated RPA32 mutations. A loading control is indicated for both gels.

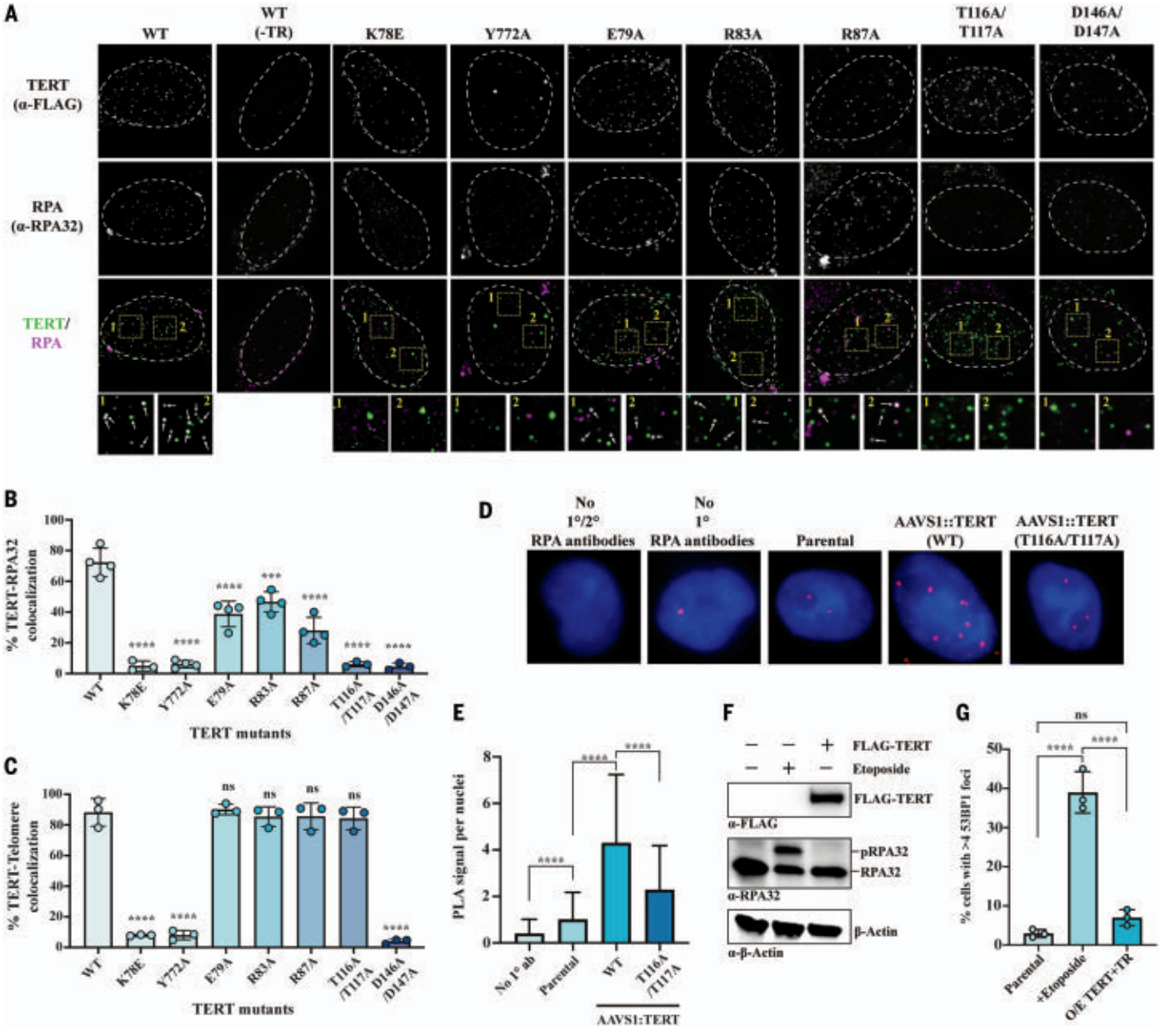


Fig. 3. RPA localizes to telomeres by means of telomerase without triggering a DDR. (**A**) Representative IF images of HeLa cells expressing WT TERT with or without TR, or the indicated TERT mutants. TERT (α-FLAG, green) and RPA32 (α-RPA32, magenta) are shown. Merged images display colocalization, with insets showing magnified views of the boxed regions where colocalization is indicated with arrows. Dashed lines indicate nuclei. (**B**) Quantification of the percentage of TERT foci that colocalize with RPA32 foci for WT and the indicated TERT mutants, based on images as in (A). Statistical significance was determined relative to the WT sample. (**C**) Quantification of the percentage of TERT foci that colocalize with telomeres for WT and the indicated TERT mutants. Statistical significance was determined relative to the WT sample. (**D**) Representative images from a PLA showing TERT–RPA32 association (magenta foci) in parental HeLa cells and in cells with WT or T116A/T117A TERT integrated at the AAVS1 locus. (**E**) Quantification of the PLA signal per nucleus from (D). (**F**) Western blot analysis of cell lysates probed for phosphorylated RPA32 (pRPA32), a marker of DDR. Etoposide treatment serves as a positive control. β-Actin is a loading control. (**G**) Quantification of the percentage of cells with 53BP1 foci, a DDR marker. The analysis compares parental cells, etoposide-treated cells, and cells overexpressing TERT and TR (O/E TERT+TR). In all graphs, error bars indicate SD. Data points in (B) and (C) are from biological replicates, with >100 total nuclei analyzed per sample. ns, not significant; ***$P \leq 0.001$; ****$P \leq 0.0001$.

stimulation (Fig. 2F). This system allows a direct probe of RPA's processivity function in cells. We used CRISPR-Cas9 to insert expression cassettes for WT or mutant TERT into the AAVS1 safe-harbor locus (*34*) in HEK293T cells (Fig. 4A). The resulting polyclonal cell lines, verified with polymerase chain reaction (PCR) and Western blot (Fig. 4B), express exogenous TERT that dominates telomerase assembly, leading to a dominant phenotype (*32*, *35*).

Southern blot analysis of telomere length revealed the functional consequences of these mutations (Fig. 4C). Whereas WT TERT expression caused telomere lengthening, several mutants led to telomere shortening. These included the recruitment-defective $TERT_{TPT}$ mutants (K78E and Y772A) and D146A/D147A mutant. The SOF $TERT_{RPA}$ mutants (R83A and T116A/T117A), which are recruited to telomeres but cannot be stimulated by RPA, also caused telomere shortening.

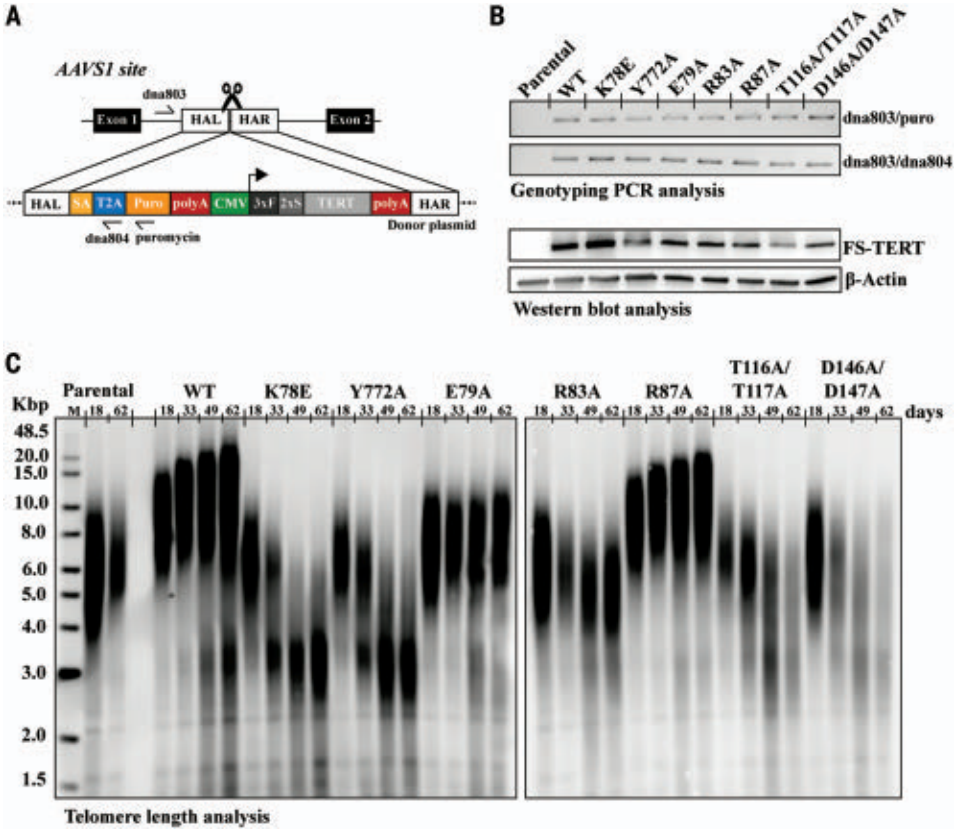


Fig. 4. TERT SOF RPA mutants are defective in telomere maintenance in cells. (**A**) Schematic of the strategy for CRISPR-Cas9–mediated insertion of a TERT expression cassette into the AAVS1 safe-harbor locus. HAL/HAR, homology arms; Puro, puromycin resistance; 2xS, Twin-Strep tag; 3xF, 3xFLAG tag. (**B**) Genotyping PCR analysis confirming successful insertion at the AAVS1 locus (top) and Western blot analysis of FLAG-Strep–tagged TERT (FS-TERT) expression in the indicated cell lines (bottom). β-Actin serves as a loading control. (**C**) Telomere length analysis by Southern blot of the indicated cell lines cultured over several time points (days). M, molecular weight marker in kilobase pairs (Kbp).

Conversely, mutants that did not disrupt stimulation by either RPA or TPT (E79A and R87A) maintained or elongated telomeres.

These findings collectively highlight the role of RPA in telomere maintenance. The key observation is that the TERT R83A and T116A/T117A mutants cannot support telomere elongation despite retaining their ability to be stimulated by TPP1–POT1. Whereas TPP1 is essential for telomerase recruitment and may contribute to processivity stimulation when partnered with POT1, RPA provides a distinctive and indispensable enhancement of RAP during telomere extension.

Connections to short telomeres human diseases

To connect our findings to telomere-related diseases, we examined several short-telomere syndrome mutations—TERT R83P (*39*), F115L (*40*), and T116I (*41*)—that map to the predicted TERT–RPA32 interface (Fig. 5A). Direct telomerase assays showed that all three mutants impaired RPA-mediated RAP stimulation (Fig. 5B and fig. S14, A and B). Whereas the R83P mutation disrupts stimulation by both RPA and TPT, likely by destabilizing the helical structure between residues 76 to 92, F115L and T116I selectively impair only RPA stimulation, which is consistent with a disruption at the RPA interface. In cells, the R83P mutant failed to localize to telomeres, which is consistent with its TPT stimulation defect (Fig. 5, C and D). By contrast, F115L and T116I localized normally, fulfilling the criteria for SOF mutants (Fig. 3C). As predicted from their impaired RPA stimulation, all three disease mutants showed a significant drop in TERT–RPA colocalization in cells (Fig. 5E and fig. S14C).

Telomere length assays in HEK293T cells revealed the functional consequences of these mutations (Fig. 5F). Unlike WT TERT, which robustly lengthened telomeres, all three disease mutants failed to support elongation. R83P caused rapid telomere shortening, which is consistent with its recruitment defect. T116I also led to shortening despite its normal localization, supporting the model that loss of RPA-mediated stimulation impairs telomerase activity in cells. The F115L mutant maintained initial telomere length but failed to promote elongation, suggesting a partial loss of function.

Recent studies have also linked short-telomere syndromes to mutations in RPA70 (*42*) and RPA32 (*43*). We investigated whether these mutations impair telomerase processivity stimulation, which would suggest additional RPA–telomerase contacts. However, direct assays showed that RPA complexes with these disease-associated variants stimulate telomerase normally (fig. S15). This suggests that the telomere defects associated with these particular mutations arise from functions distinct from the direct telomerase stimulation we described here.

Discussion

Our work demonstrates that the human RPA heterotrimer directly enhances telomerase RAP. This stimulation requires RPA to bind both the TERT TEN domain and the 3′ telomeric overhang and that disrupting this interaction causes telomere shortening in our cellular assays. We propose a sequence in which telomerase is first recruited to telomeres by TPP1 (*10–12*, *44–46*) and subsequently engaged by RPA for processive extension of the telomere overhang. Whether TPP1–POT1 contributes to processivity in cells remains to be determined. The inactivation of DDR at telomeres despite the presence of RPA can be explained by POT1 capping the telomere duplex–ssDNA junction (*47*), preventing the activation of RPA by ATR.

The idea that an RPA-family complex stimulates telomerase was foreshadowed by findings in other systems (*15–19*). In *Tetrahymena*, Teb2 and Teb3 are the RPA32 and RPA14 subunits, respectively, and together with Teb1 form the TEB complex (*18*). Teb2 binds the same TERT surface predicted here for RPA32 OB-D, and an intact TEB complex is required for maximal telomerase stimulation activity (*17*). While this manuscript was in revision, a genetic study in budding yeast suggested that RPA interacts with the telomerase subunit Est3 in vivo and stimulates activity (*19*). Taken together with our results, these observations suggest that RPA-family complexes play an evolutionarily conserved role as direct modulators of telomerase RAP.

Comparison of our TERT–RPA32 AlphaFold model with the telomerase–TPP1–POT1 cryo-EM structure (*29*) suggests a potential steric clash between POT1 and RPA32 because they bind overlapping surfaces on TERT (fig. S16). This implies that the initial TPP1–POT1 complex may need to be displaced for RPA to engage telomerase. To explore this possibility, we used AlphaFold3 (*31*) to model the complete assembly. AlphaFold3 predicted a "hand-over" conformation that repositioned POT1, with RPA70 instead binding the 5′ overhang, while TPP1 and RPA32 simultaneously engage the TERT TEN domain (fig. S5).

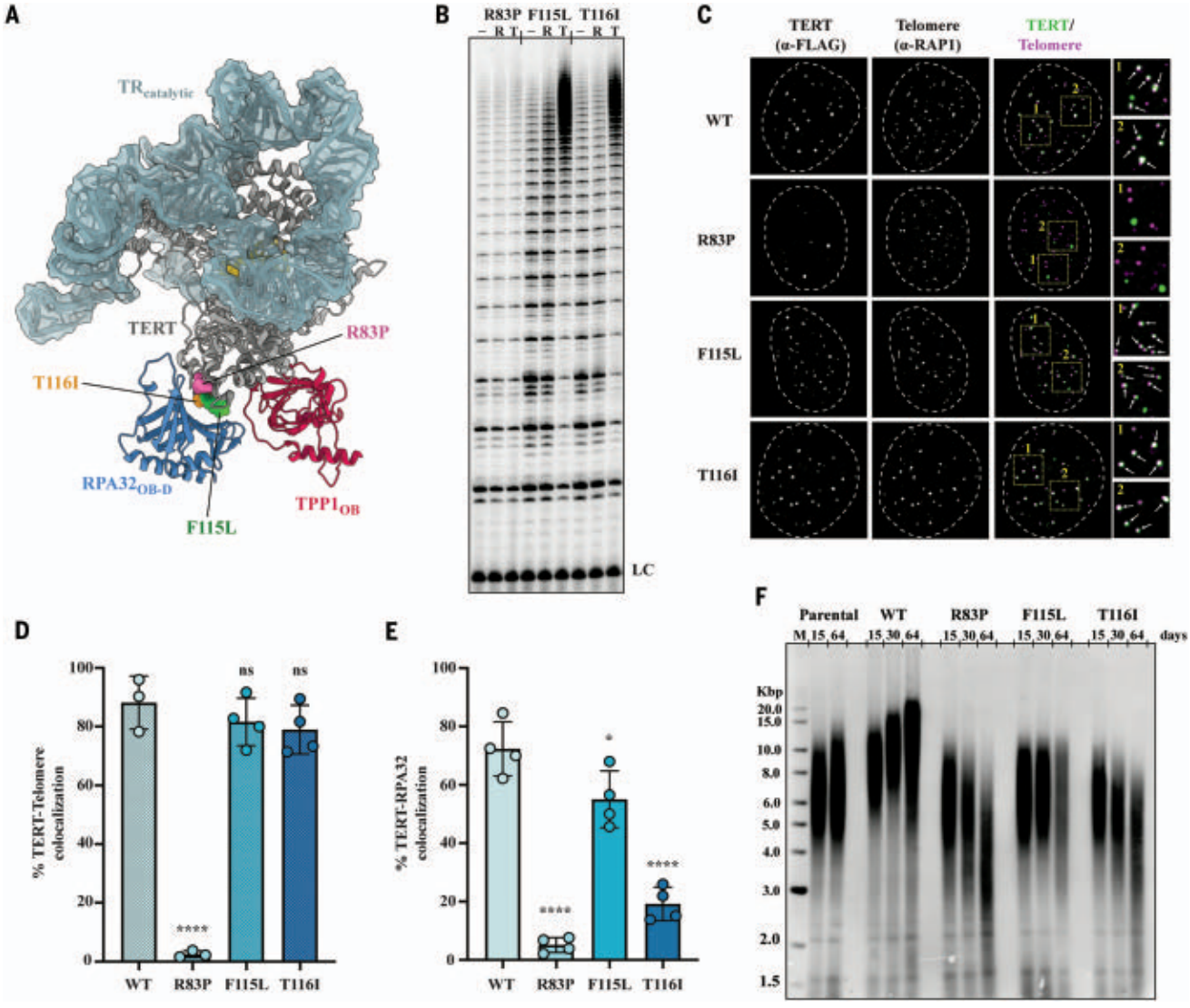


Fig. 5. Loss in telomerase RAP stimulation by RPA provides an explanation for short telomere disease TERT mutations. (**A**) Short-telomere disease–associated TERT mutations mapped onto the structural model of the telomerase-RPA32-TPP1 complex. R83P (pink) (*39*), F115L (green) (*40*), and T116I (orange) (*41*) are highlighted. (**B**) Direct telomerase assay comparing WT TERT with the indicated disease mutants in the presence of no added protein (–), 200 nM RPA (R), or 200 nM TPT (T). (**C**) Representative IF images showing localization of TERT (α-FLAG, green) and telomeres (α-RAP1, magenta) in cells expressing WT TERT or the indicated mutants. Dashed lines indicate nuclei. (**D**) Quantification of TERT and telomere colocalization from images as in (C). (**E**) Quantification of TERT and RPA32 colocalization for the indicated mutants. (**F**) Telomere length analysis by means of Southern blot of parental cells and cells expressing WT or mutant TERT at different time points (days). M, molecular weight marker in kilobase pairs (Kbp). In bar charts, error bars indicate the SD from $n = 3$ independent experiments, with >100 nuclei analyzed per condition for IF quantifications in (D) and (E). ns, not significant; $^{*}P \leq 0.05$; $^{****}P \leq 0.0001$.

Our biochemical data are consistent with this proposed hand-over model. The AlphaFold3 model predicts that RPA70 OB-A and -B domains bind the 5′ end of the telomeric overhang, which agrees with our finding that these ssDNA binding domains are essential for stimulation (Fig. 1E). Furthermore, RPA displaced a prebound TPP1–POT1 complex from ssDNA regardless of POT1's binding position (fig. S17, A and B), which is consistent with a previous report (*48*). Adding both factors together did not further increase processivity, suggesting that only one factor is used at a time (fig. S17, C to E). Collectively, these results support a two-step mechanism in which RPA replaces POT1 to engage the overhang and stimulate telomerase processivity. Although this model fits our biochemical data and AlphaFold predictions, alternative coordination mechanisms involving POT1—with or without TPP1—remain possible. Future studies will be necessary to clarify the precise sequence of events and the potential roles of other telomere-associated factors (*48*).

Our findings have implications for understanding neo-telomere formation and telomere-related diseases. Because RPA is a major component at double-stranded DNA breaks, our discovery that it stimulates telomerase processivity raises the possibility that this mechanism facilitates telomere formation at these sites (*49*, *50*). Furthermore, identifying disease-associated TERT mutations that disrupt this RPA-mediated stimulation provides a direct mechanistic link to short-telomere syndromes and highlights the RPA–telomerase axis as a potential therapeutic target. Future studies should delineate how shelterin, RPA, and telomerase coordinate to regulate telomere maintenance.

REFERENCES AND NOTES

1. R. J. O'Sullivan, J. Karlseder, *Nat. Rev. Mol. Cell Biol.* **11**, 171–181 (2010).
2. C. Lim, T. R. Cech, *Nat. Rev. Mol. Cell Biol.* **22**, 283–298 (2021).
3. H. Takai, V. Aria, P. Borges, J. T. P. Yeeles, T. de Lange, *Nature* **627**, 664–670 (2024).
4. C. W. Greider, E. H. Blackburn, *Cell* **51**, 887–898 (1987).
5. C. W. Greider, E. H. Blackburn, *Nature* **337**, 331–337 (1989).
6. J. Lingner *et al.*, *Science* **276**, 561–567 (1997).
7. T. M. Nakamura *et al.*, *Science* **277**, 955–959 (1997).
8. Q. He, C. Lim, *Trends Biochem. Sci.* **48**, 860–872 (2023).
9. S. W. Cai, T. de Lange, *Genes Dev.* **37**, 555–569 (2023).
10. H. Xin *et al.*, *Nature* **445**, 559–562 (2007).
11. E. Abreu *et al.*, *Mol. Cell. Biol.* **30**, 2971–2982 (2010).
12. A. M. Tejera *et al.*, *Dev. Cell* **18**, 775–789 (2010).
13. F. Wang *et al.*, *Nature* **445**, 506–510 (2007).
14. M. S. Wold, *Annu. Rev. Biochem.* **66**, 61–92 (1997).
15. V. Schramke *et al.*, *Nat. Genet.* **36**, 46–54 (2004).
16. P. Luciano *et al.*, *EMBO J.* **31**, 2034–2046 (2012).
17. J. Jiang *et al.*, *Science* **350**, aab4070 (2015).
18. H. E. Upton, H. Chan, J. Feigon, K. Collins, *J. Biol. Chem.* **292**, 217–228 (2017).
19. C. A. Moeller-McCoy *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2419309122 (2025).
20. M. P. Rubtsova *et al.*, *Biochemistry (Mosc.)* **74**, 92–96 (2009).
21. A. Bochkarev, R. A. Pfuetzner, A. M. Edwards, L. Frappier, *Nature* **385**, 176–181 (1997).
22. P. Baumann, T. R. Cech, *Science* **292**, 1171–1175 (2001).
23. D. Loayza, H. Parsons, J. Donigian, K. Hoke, T. de Lange, *J. Biol. Chem.* **279**, 13241–13248 (2004).
24. C. M. Latrick, T. R. Cech, *EMBO J.* **29**, 924–933 (2010).
25. J. Jumper *et al.*, *Nature* **596**, 583–589 (2021).
26. L. I. Toledo *et al.*, *Cell* **155**, 1088–1103 (2013).
27. G. E. Ghanim *et al.*, *Nature* **593**, 449–453 (2021).
28. B. Liu *et al.*, *Nature* **604**, 578–583 (2022).
29. Z. Sekne, G. E. Ghanim, A.-M. M. van Roon, T. H. D. Nguyen, *Science* **375**, 1173–1176 (2022).
30. S. Padmanaban, V. M. Tesmer, J. Nandakumar, *Life Sci. Alliance* **6**, e202201727 (2023).
31. J. Abramson *et al.*, *Nature* **630**, 493–500 (2024).
32. J. C. Schmidt, A. B. Dalby, T. R. Cech, *eLife* **3**, e03563 (2014).
33. O. Söderberg *et al.*, *Nat. Methods* **3**, 995–1000 (2006).
34. R. C. DeKelver *et al.*, *Genome Res.* **20**, 1133–1142 (2010).
35. L. Xi, T. R. Cech, *Nucleic Acids Res.* **42**, 8565–8577 (2014).
36. T. L. Beattie, W. Zhou, M. O. Robinson, L. Harrington, *Mol. Cell. Biol.* **21**, 6151–6160 (2001).
37. E. L. Denchi, T. de Lange, *Nature* **448**, 1068–1071 (2007).
38. S. Din, S. J. Brill, M. P. Fairman, B. Stillman, *Genes Dev.* **4**, 968–977 (1990).
39. T. J. Vulliamy *et al.*, *PLOS ONE* **6**, e24383 (2011).
40. K. E. Schratz *et al.*, *Blood* **135**, 1946–1956 (2020).
41. S. Feurstein *et al.*, *Blood Adv.* **4**, 4873–4886 (2020).
42. R. Sharma *et al.*, *Blood* **139**, 1039–1051 (2022).
43. R. Kochman *et al.*, *Genes Dev.* **38**, 755–771 (2024).
44. F. L. Zhong *et al.*, *Cell* **150**, 481–494 (2012).
45. J. Nandakumar *et al.*, *Nature* **492**, 285–289 (2012).
46. A. N. Sexton *et al.*, *Genes Dev.* **28**, 1885–1899 (2014).
47. V. M. Tesmer, K. A. Brenner, J. Nandakumar, *Science* **381**, 771–778 (2023).
48. R. L. Flynn *et al.*, *Nature* **471**, 532–536 (2011).
49. A. O. M. Wilkie, J. Lamb, P. C. Harris, R. D. Finney, D. R. Higgs, *Nature* **346**, 868–871 (1990).
50. C. G. Kinzig, G. Zakusilo, K. K. Takai, L. R. Myler, T. de Lange, *Science* **383**, 763–770 (2024).
51. M. A. Cianfrocco, M. Wong-Barnum, C. Youn, R. Wagner, A. Leschziner, "COSMIC2: A Science Gateway for Cryo-Electron Microscopy Structure Determination" in *Practice and Experience in Advanced Research Computing 2017: Sustainability, Success and Impact* (Association for Computing Machinery, 2017), pp. 1–5; https://dl.acm.org/doi/10.1145/3093338.3093390.
52. S. Agrawal *et al.*, Human RPA is an essential telomerase processivity factor for maintaining telomeres. Zenodo (2025); https://doi.org/10.5281/zenodo.17100911.

ACKNOWLEDGMENTS

We thank V. Lundblad, D. Wuttke, M. Stone, J. Schmidt, S. Cohen, and T. Bryan for their constructive suggestions and discussions and T. Record and colleagues in the Lim and Huang laboratories for their valuable input. We thank K. Nguyen for generously providing the hTR expression plasmid and K. Mehta for the superRPA plasmid. For computational support, we acknowledge the COSMIC^2 platform (*51*) for our initial AlphaFold work and thank J. Herr and the departmental IT team for their assistance in setting up our computational workflow. We acknowledge the use of facilities and equipment supported by the University of Wisconsin–Madison Biophysics Instrumentation Facility, which was established with support from the University of Wisconsin–Madison and grants BIR-9512577 (NSF) and S10 RR13790 (NIH). **Funding:** Support for this research was provided to C.J.L. by the National Institutes of Health; the National Institute of General Medical Sciences (R01GM153806 and DP2GM150023); and the University of Wisconsin–Madison, Office of the Vice-Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation and the Department of Biochemistry. X.H. acknowledges the support from the Hirschfelder Professorship Fund. In addition, individual support was provided to M.S.O. and K.M.A. by a NIH T32 predoctoral fellowship (T32GM130550), to S.A. by the William H. Peterson Graduate Fellowship, to B.L.C. by the Arthur B. Michael Biochemistry Graduate Fellowship, and to Q.H. by the Louis and Elsa Thomsen Wisconsin Distinguished Graduate Fellowship. **Author contributions:** Conceptualization: C.J.L. Methodology: S.A., X.L., V.S., M.S.O., B.L.C., V.R.T., X.H., and C.J.L. Investigation: S.A., X.L., V.S., M.S.O., B.L.C., G.P.T., V.R.T., Q.H., K.M.A., and C.J.L. Visualization: V.S. and C.J.L. Funding acquisition: C.J.L. and X.H. Project administration: C.J.L. Supervision: S.A., C.J.L., and X.H. Writing – original draft: C.J.L. and S.A. Writing – review & editing: C.J.L, S.A., X.L., V.S., M.S.O., B.L.C., V.R.T., Q.H., K.M.A., and X.H. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Plasmids and cell lines generated are available upon request. All AlphaFold predicted models and MD simulation–related files are archived at Zenodo (*52*). All other data are available in the main text or supplementary materials. **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ads5297
Materials and Methods; Figs. S1 to S17; Tables S1 to S3; References (*53*–*76*)

Submitted 17 August 2024; resubmitted 8 June 2025; accepted 11 September 2025

10.1126/science.ads5297

POLYMER RECYCLING

Glycolysis-compatible urethanases for polyurethane recycling

Yanchun Chen[1]†, Jinyuan Sun[1,2]†, Kelun Shi[1,2]†, Tong Zhu[1], Ruifeng Li[1], Ruiqiao Li[2,3], Xiaomeng Liu[1,2], Xinying Xie[1], Chao Ding[1], Wen-Chao Geng[1,4], Jinwei Ren[5], Wenyu Shi[6], Yinglu Cui[1,5,7]*, Bian Wu[1]*

Recycling thermoset polyurethanes is hindered by their cross-linked structures and chemically stable urethane bonds. Although chemo-enzymatic approaches offer promise, known urethanases remain inefficient under industrial glycolysis conditions. Here, we present GRASE [graph neural network (GNN)–based recommendation of active and stable enzymes], a GNN-based framework that integrates self-supervised and supervised learning to identify efficient, glycolysis-compatible urethanases. Among these, *Ab*PURase exhibited two orders of magnitude greater activity than previously known enzymes in 6 molar diethylene glycol, enabling near-complete depolymerization of commercial polyurethane at kilogram scale within 8 hours. Structural analysis revealed that a tightly packed hydrophobic core and proline-stabilized lid loop may confer *Ab*PURase's stability and efficiency in harsh solvents. This work highlights how deep learning accelerates the discovery of biocatalysts with industrial potential and addresses a critical barrier in polyurethane recycling.

Advances in biocatalytic polyester recycling (*1*) have illuminated the potential of enzymatic depolymerization for developing alternatives, especially hydrolyzable polymers. Polyurethane, with an annual global consumption of 22 million metric tons in 2024, constitutes the second largest category of hydrolyzable plastics (*2*, *3*). Its production process is energy intensive and heavy in terms of greenhouse gas emissions, requiring 331 MJ/kg of energy and 12.3 kg-CO_2e/kg of greenhouse gas emissions (*4*). Polyurethanes are classified as polyester or polyether based depending on the polyol segment, although all contain urethane linkages. Recent biodegradation efforts have focused on ester hydrolases such as cutinases, which cleave ester bonds in the polyol segment (*5*–*9*). By contrast, polyether polyurethanes, which constitute ~75% of the polyurethane foam market (*10*), remain highly resistant to biodegradation because of the lack of enzymes that are capable of cleaving urethane linkages (*11*).

Compounding this challenge, commercial polyurethane materials are predominantly thermoset with extensive cross-linked networks that impede enzymatic penetration. Unlike thermoplastics such as polyethylene terephthalate (PET), which can be remolded into amorphous states through melt extrusion, thermoset polyurethanes cannot be reshaped to enhance susceptibility to enzymatic attack (*12*, *13*). Therefore, alternative recycling techniques are needed to overcome the depolymerization challenges posed by thermoset materials.

[1]AIM Center, College of Life Sciences and Technology, Beijing University of Chemical Technology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]State Key Laboratory of Biochemical Engineering, Institute of Process Engineering, Chinese Academy of Sciences, Beijing, China. [4]State Key Laboratory of Green Biomanufacturing, Beijing, China. [5]State Key Laboratory of Microbial Diversity and Innovative Utilization, Beijing, China. [6]College of Biological Sciences, China Agricultural University, Beijing, China. [7]Beijing Key Laboratory of Genetic Element Biosourcing & Intelligent Design for Biomanufacturing, Beijing, China. ***Corresponding author. Email: cuiyinglu@im.ac.cn (Y.Cui); thebianwu@outlook.com (B.W.)** †These authors contributed equally to this work.

Currently, glycolysis represents the most promising industrial-scale chemical recycling method for polyurethane. During this process, polyols undergo transesterification with alcohol reagents such as diethylene glycol (DEG) (*14*–*16*), resulting in an upper phase enriched with recovered polyols that are suitable for straightforward reuse in new polyurethane syntheses. The bottom phase, constituting >60% (w/w) of glycolysis products, accumulates excess DEG along with a mixture of *N*-aryl carbamates [toluenediamine-DEG (TDA-DEG) or 4,4′-methylenedianiline-DEG (MDA-DEG)] and their related derivatives. Although the polyol-rich upper phase is routinely reclaimed, the bottom phase is seldom valorized and would be incinerated as hazardous waste, imposing environmental and economic burdens. The aromatic mixture can increase the viscosity of the bottom phase, which severely impedes the DEG recovery, so ongoing research has focused on valorizing the aromatic mixture (*11*, *17*). Although hydrolytic recovery offers a potential route for recovering monomeric components from these carbamate by-products, the process currently requires energy-intensive conditions (200°C, 16 bar, pH > 12), diminishing the ecological advantages provided by glycolysis (*17*) (fig. S1, tables S1 and S2, and supplementary text 1). A recent advancement by Branson *et al.* introduced a chemo-enzymatic process for polyurethane foam recycling using urethanases (UMG-SPs, including SP1, SP2, and SP3) to hydrolyze *N*-aryl carbamates (*11*). Despite considerable interest, these enzymes retain unresolved limitations in organic solvent compatibility, a longstanding challenge in chemo-enzymatic systems (*18*). A phase change from the organic phase to the aqueous phase is necessary for facilitating enzymatic depolymerization, which requires a substantial amount of solvent evaporation that compromises industrial feasibility (table S2 and supplementary text 1). Further protein engineering efforts to enhance the performance of SP1 have yielded limited gains (*19*), raising demands for more efficient glycolysis-compatible enzymes.

Distinct regioselectivity across diverse urethanases

With various potential PURases described, we initially explored the hydrolytic activity of 14 literature-claimed PURases on *N*-aryl carbamates, encompassing carbamate-degrading amidases from *Agrobacterium tumefaciens* d3 (AmdA) (*20*), *Rhodococcus erythropolis* MP_{50} (ReMP50) (*21*), *Candida parapsilosis* NBRC 708 (CPUTNase) (*22*), *Aspergillus oryzae* (AmdS) (*22*), and *Lysinibacillus fusiformis* SCO2 (SCO2) (*23*), as well as SP1 (*11*) and seven metagenomic urethanases from a patent (GatA or Aes urethanases) (*24*). LCC^{ICCG} (*1*), a cutinase known for its activity on PET, was also evaluated. For benchmarking, we used aromatic carbamates derived from the glycolysis of toluene diisocyanate (TDI)–based polyurethane with DEG. The bottom-phase products included bis[2-(2-hydroxyethoxy)ethyl] (4-methyl-1,3-phenylene) dicarbamate (2,4-TDA-DEG) and bis[2-(2-hydroxyethoxy)ethyl] (2-methyl-1,3-phenylene) dicarbamate (2,6-TDA-DEG), along with their derivatives. We mapped the *ortho*- versus *para*-hydrolytic specificities of each enzyme given the enzymatic accessibility of either ester group in these carbamates.

Among the 14 enzymes evaluated, only Aes72 (*24*), GatA (*24*), and SP1 demonstrated hydrolytic activity toward these aromatic carbamates, with Aes72 exhibiting the highest specific activity of 0.06 U/mg for the hydrolysis of 2,4-TDA-DEG (Fig. 1). This enzyme preferentially hydrolyzed the *ortho* urethane bond of 2,4-TDA-DEG to generate the intermediate 4-TAC, which subsequently underwent hydrolysis at the *para* position to produce the 2,4-TDA monomer. Although GatA also targeted the *ortho* urethane bond, it exhibited considerably lower efficiency (<60% substrate conversion) and failed to produce any detectable monomers within the reaction period. In contrast to Aes72 and GatA, SP1 demonstrated more promiscuous regioselectivity, hydrolyzing both *ortho* and *para* urethane bonds. Nevertheless, SP1 primarily generated 2-TAC as the intermediate of 2,4-TDA-DEG with low catalytic efficiency, whereas it showed higher activity (0.65 U/mg) toward

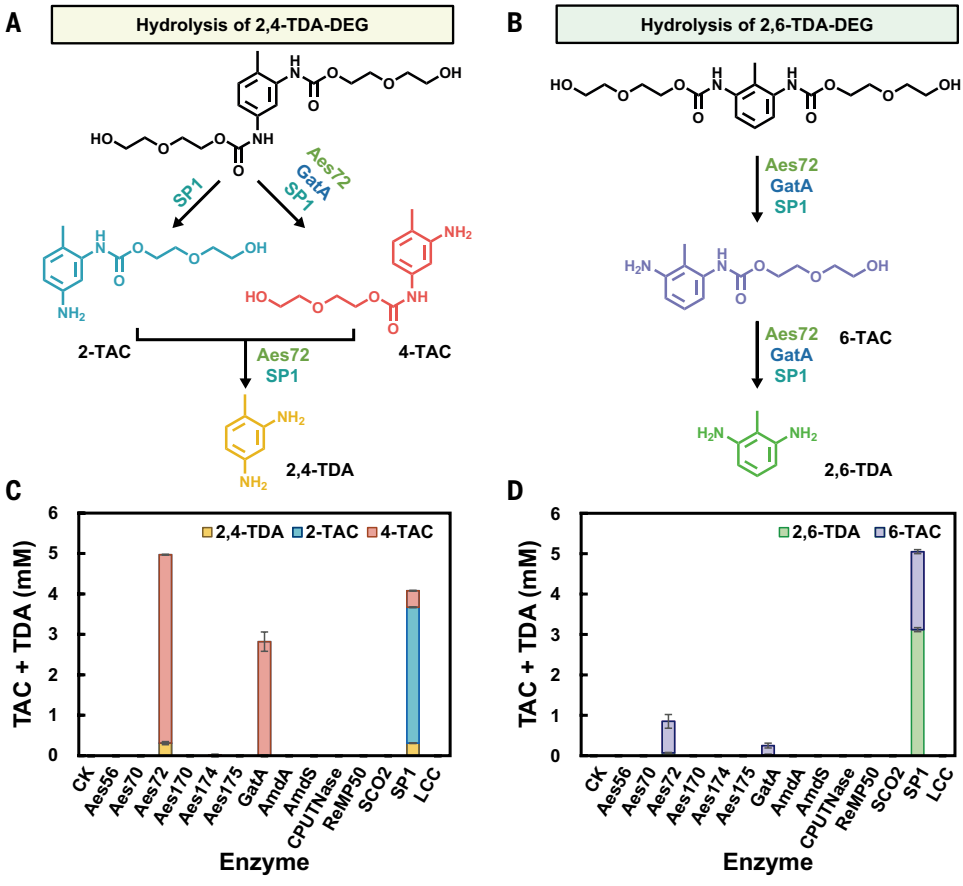


Fig. 1. Hydrolytic pathway of TDA-DEG and the enzymatic activity of literature-claimed PURases. (**A**) The hydrolysis of 2,4-TDA-DEG produces the intermediate compounds 2-TAC and 4-TAC, which are further hydrolyzed to produce 2,4-TDA. (**B**) The hydrolysis of 2,6-TDA-DEG generates the intermediate 6-TAC, which is subsequently hydrolyzed to form 2,6-TDA. (**C** and **D**) Enzymatic hydrolysis of 2,4-TDA-DEG (C) and 2,6-TDA-DEG (D) through published PURases. Reactions were conducted in 100 mM sodium phosphate buffer, pH 7.5, at 50°C for all enzymes except SP1, which was performed at 37°C to prevent thermal inactivation. All data are presented as mean ± SD ($n = 3$).

2,6-TDA-DEG (~44-fold higher activity than that of Aes72). Given the typical 8:2 predominance of 2,4-TDA-DEG over 2,6-TDA-DEG in feedstocks (*25*), preferentially hydrolyzing of 2,4-TDA-DEG is more essential. Therefore, Aes72 has emerged as a promising template for further development.

Machine learning–guided discovery of efficient PURases

We commenced our exploration by performing sequence-based searches with BLASTp against the National Center for Biotechnology Information (NCBI) nonredundant protein database, as well as structure-based retrieval with Foldseek (*26*) against the AlphaFold Protein Structure Database (AlphaFold DB). Candidates identified by these approaches were prioritized on the basis of their sequence or structural identity with respect to Aes72. An experimental validation of the top 10 ranked candidates derived from these approaches demonstrated that only four exhibited urethanase activity for generating TDA monomers, with candidate B1 showing a modest 30% increase in activity compared with that of Aes72 (table S3). Candidates F4 and F5 even failed to generate either TAC intermediates or TDA monomers. We subsequently adopted a clustering approach analogous to that used for exploring the PETase sequence landscape (*27*) and complemented it with an enzyme-substrate complex structure prediction method for substrate-binding analysis (*28*). A total of 216 representative sequences were sampled from 126 clusters through a neighborhood analysis, and the structures were predicted for each in complex with 2,4-TDA-DEG. The 12 top-ranked candidates with the lowest Rosetta energy scores were selected for experimental validation. However, only two candidates generated TDA monomers (table S4). These results show the limitations of relying exclusively on sequence or binding predictions for PURase discovery without extensive experimental characterization.

Recently, deep learning models that learn high-dimensional enzyme representations have demonstrated greater predictive accuracy than that of conventional sequence-similarity metrics, particularly for enzyme commission (EC) number prediction tasks (*29*). Given the absence of specific EC numbers for PURases, we introduced a graph neural network (GNN)–based algorithm using supervised learning frameworks to capture structural representations for function prediction. The pocket residue prediction neural network (named Pythia-Pocket) was trained on the PDB-Bind database to extract the functional features of individual amino acids from protein structures and to predict the probability of each amino acid being a ligand-binding residue (Fig. 2A). Pythia-Pocket used predicted apo structures as its inputs without relying on the additional enzymatic activity data required by other pocket-based approaches (*30*). Protein structures were abstracted as graphs, where amino acid types were encoded as vertex features and backbone atomic distances were encoded as edge features. Neighboring residues were defined using the k-nearest neighbor approach. A message-passing neural network (MPNN) was used to iteratively update the vertex and edge representations. Through multiple MPNN blocks, information propagated across the structure, ultimately generating residue-level embeddings. The embeddings were then compared using cosine similarity to assess their structural similarity. Higher similarity among the catalytic residue representations indicated analogous pocket environments, thereby facilitating the identification of functionally related enzymes. To evaluate the efficiency of Pythia-Pocket, we benchmarked it against the existing PETase identification methods. The results revealed the outstanding performance of Pythia-Pocket in identifying experimentally validated PETases, highlighting the advantage of this approach in pinpointing functionally similar enzymes, especially within underrepresented enzyme classes (fig. S2, table S5, and supplementary text 2.1 and 2.2). The stability of the candidates was evaluated by Pythia (*31*), our previously developed neural network, which predicts the likelihood of an individual residue on the base of its local structural context, to estimate protein stability through a negative log-likelihood (NLL) metric (fig. S3 and supplementary text 2.3 and 2.4). In light of the inherent trade-off between activity and stability, we integrated both Pythia and Pythia-Pocket into a unified framework called GRASE (GNN-based recommendation of active and stable enzymes) to identify candidates that optimally balanced the structural robustness and target functionality of the NCBI nonredundant protein database and the AlphaFold DB (Fig. 2B). We stratified the cosine similarity between

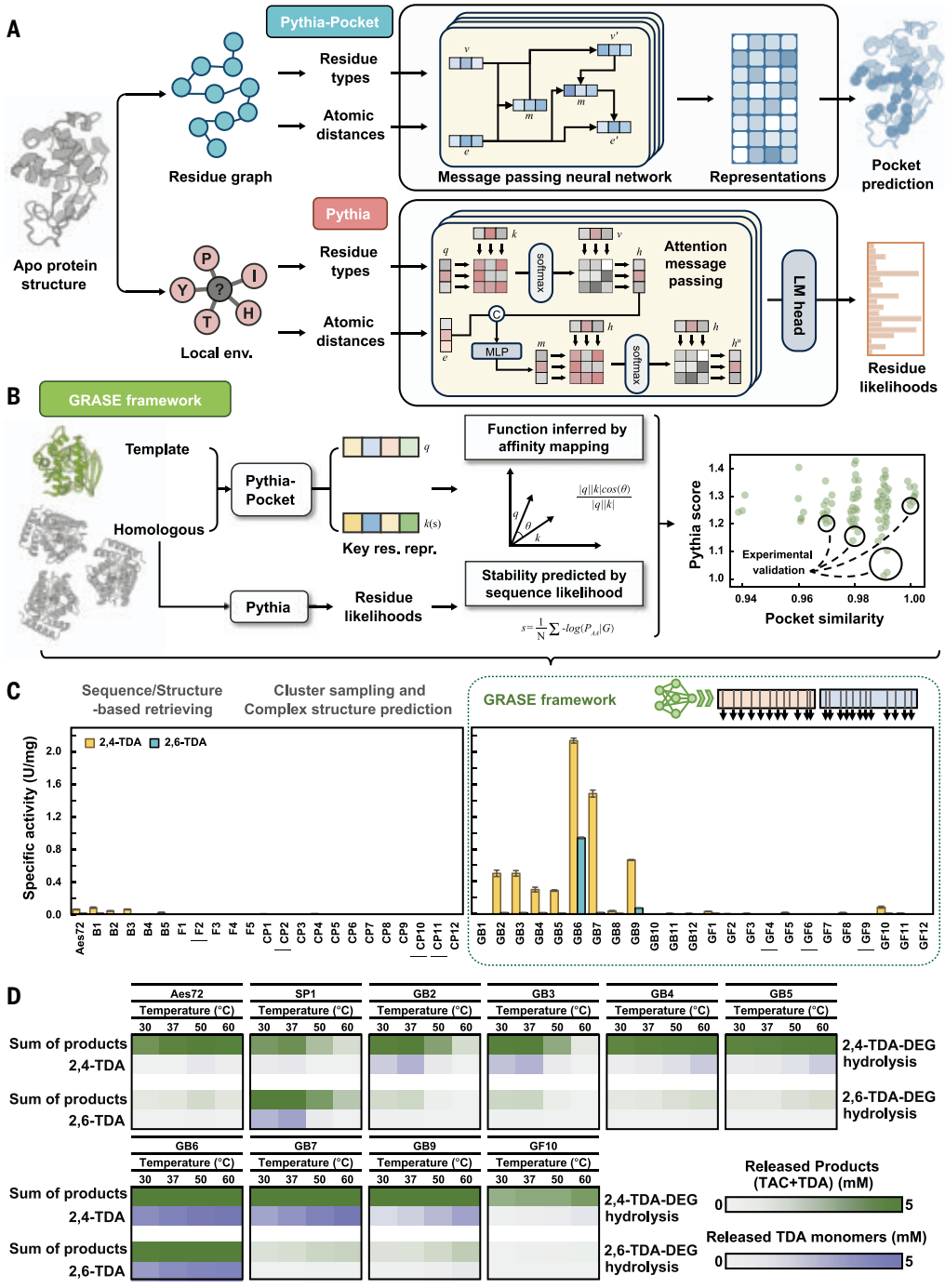


Fig. 2. Architecture of the models and functional validation of GRASE-identified enzymes. (**A**) Architectures of Pythia-Pocket and Pythia. Both models abstract protein structures as graphs, starting with the apo protein structure. In Pythia-Pocket, the entire protein graph is used as the input for an MPNN, generating vertex representations (v) for each residue. These representations are then passed through a linear layer to predict the probability of each residue being part of the ligand-binding pocket. Conversely, Pythia uses the local structure context of each amino acid as an input, updating residue-specific interactions through attention-based message passing. The final output predicts probability distributions for each amino acid type at the central residue. (**B**) Overview of the GRASE workflow. Pythia-Pocket generates residue-level embeddings for the key catalytic residues of the query protein (e.g., Aes72) and candidate structures. The cosine similarity between the key residue representations is calculated to quantify the functional similarity between the active sites of the query and candidate proteins. Concurrently, protein stability is evaluated using Pythia by calculating the NLL of each candidate sequence, where lower NLL values indicate higher predicted stability. (**C**) Enzyme hydrolysis of 2,4-TDA-DEG and 2,6-TDA-DEG. The candidates identified through a sequence-based search (B1 to B5), structure-based retrieving (F1 to F5), and sequence clustering (CP1 to CP12), along with GRASE-curated enzymes (GB1 to GB12 and GF1 to GF12) were evaluated for their catalytic activity toward 2,4-TDA-DEG and 2,6-TDA-DEG. Enzymes that failed to express (F2, CP2, CP10, CP11, GF4, GF6, and GF9) are underlined. The specific activities of these enzymes were measured at their optimal temperatures. One unit of PURase activity is defined as the amount of enzyme that hydrolyze 2,4-TDA-DEG equivalent to 1.0 μmol of 2,4-TDA monomers per minute. All data are presented as mean ± SD ($n = 3$). (**D**) Heatmaps of each product after TDA-DEG hydrolysis with GRASE-curated enzymes performed from 30° to 60°C over 1 hour in 100 mM sodium phosphate buffer, pH 7.5. Enzymes with specific activities lower than 0.05 U/mg are not shown. The heatmap gradient represents the sum of the released products (TAC intermediates + TDA monomers) (green) and TDA monomers (purple) from 0 to 5 mM.

the key catalytic residues of the candidates and Aes72 derived from Pythia-Pocket into four distinct tiers ranging from 1.00 to 0.97. Within each tier, the three sequences with the highest Pythia-predicted stability values were selected. This approach established functional similarity as a baseline while maximizing stability, allowing for a more comprehensive exploration of the sequence-function landscape.

By applying GRASE without any expert curation, we experimentally evaluated the 24 top-ranked candidates (tables S6 and S7). Among them, 21 were successfully expressed and exhibited urethanase activity toward *N*-aryl carbamates, with 17 generating both TAC intermediates and TDA monomers and the remaining four yielding only TAC (Fig. 2, C and D). Eight outperformed the benchmark (Aes72), with five (GB2, GB3, GB6, GB7, and GB9) achieving activities greater than 0.5 U/mg (table S8). GB6 (*Ab*PURase) exhibited the highest activity, with a 32-fold increase in activity against 2,4-TDA-DEG (2.14 U/mg) and a 62-fold enhancement against 2,6-TDA-DEG (0.94 U/mg) relative to Aes72. A similar performance was observed for GB7, which demonstrated activity toward 2,4-TDA-DEG at 1.49 U/mg but significantly reduced activity toward 2,6-TDA-DEG at only 0.02 U/mg. When exposed to 6 M DEG, the activity of *Ab*PURase was even more pronounced, with 465- and 170-fold improvements over Aes72 at 50° and 37°C, respectively (figs. S4 and S5). Coupled with its high expression level, these results highlighted the considerable potential of *Ab*PURase for degradation in the postglycolysis phases of polyurethane recycling. Beyond the identification of *Ab*PURase, the enzymes filtered by GRASE outperformed those ranked by overall structural similarity, because all soluble GRASE candidates produced TAC, and seven of them could further produce TDA monomers. These enzymes retained the key catalytic functions despite their low overall sequence identity levels (as low as 34%). These results illustrate that the similarity of local structural motifs may preserve the essential catalytic functions in distantly related enzymes, offering valuable perspectives on functional domains that may inform evolutionary diversification.

Chemo-enzymatic depolymerization of commercial polyurethane foams at kilogram scale

We subsequently conducted chemo-enzymatic depolymerization on commercial polyurethane foams. Given that BASF is one of the major global suppliers of polyurethane (*32*), its products are well suited for use as standardized substrates for chemo-enzymatic degradation. We embarked on the glycolysis of a TDI-based flexible polyurethane foam from BASF, resulting in an apolar upper phase enriched with polyols, an intermediate phase concentrated in styrene-acrylonitrile (*33*), and a polar bottom phase containing DEG and diverse aromatic compounds (Fig. 3A). Because of the water added during the foaming process, the polyurethane foam contained not only carbamates but also *N*,*N*′-diarylurea functional groups, leading to the formation of a small amount of diphenylurea derivatives in the bottom phase (fig. S6).

Before enzymatic treatment, the reaction parameters were optimized (figs. S7 and S8) with consideration of process-modeled operating costs for polyurethane recycling (tables S9 and S10 and supplementary text 1). An enzyme loading of 8 mg/g (milligrams of enzyme per gram of polyurethane) was estimated as the economical compromise between productivity and cost, reducing the incurred operating expenses by \$47 tons^{-1} compared with those required at 16 mg/g while achieving near-complete conversion within 12 hours (fig. S7 and table S10). After optimization, we evaluated the hydrolytic performance of *Ab*PURase, Aes72, and SP1 on the bottom phase derived from glycolysis of 30 g of BASF foam at temperatures ranging from 30° to 60°C (fig. S9 and tables S11 and S12). *Ab*PURase showed a maximal initial conversion (37% at 0.5 hours) at 60°C, but enzyme inactivation curtailed the overall depolymerization to 91% after 12 hours. At 50°C, *Ab*PURase maintained robust activity, obtaining 33% conversion at 0.5 hours and, ultimately, 98% conversion at 12 hours, with the complete hydrolysis of TAC to TDA, releasing ~385 mM of TDA. By contrast, SP1 and Aes72 achieved only 3 and 5% conversion within the initial 0.5 hours, respectively, at their optimal temperatures (30°C for SP1 and 37°C for Aes72). Prolonged 12-hour incubations led to a modest final conversion of 15% for SP1 and 26% for Aes72, rendering them unsuitable for integration into upstream glycolysis workflows (Fig. 3B). In addition to its operational stability, *Ab*PURase demonstrated notable reusability. Without any immobilization treatment, the enzyme could be recycled and achieve near-complete conversion within 12 hours for at least three cycles (fig. S10).

Building on these bench-scale results, we implemented a kilogram-scale process. One kilogram of BASF polyurethane foam was subjected to glycolysis at 190°C for 3.5 hours with DEG at a ratio of 1:1.5. This reaction afforded an upper phase of 734 g, a bottom phase of 1507 g, and 221 g of insoluble solids (Fig. 3A). Because of the excess of alcohol reagent, the bottom phase was primarily composed of 83% DEG at 8.6 M, along with 14% aromatic compounds (Fig. 3A, fig. S11, and table S13). Consistent with the bench-scale results, *Ab*PURase reached 90% conversion within 2 hours and 98.6% after 8 hours at an enzyme loading of 16 mg/g (Fig. 3C). Reducing this to 8 mg/g for economic considerations resulted in slightly diminished initial efficiency (72% at 2 hours, 90% at 4 hours, and 95% at 8 hours). The reaction progressed more gradually after 4 hours and finally reached 98% conversion after 12 hours. The resulting TDA monomers were efficiently recovered through rectification, achieving TDA and DEG recovery rates of 94.7 and 98.5%, respectively (table S14). On the basis of these results and supporting process modeling, preliminary estimates suggest promising scalability. Nevertheless, further efforts are required to improve enzyme performance through protein engineering, to optimize the conversion of recovered TDA to TDI, and to investigate the repolymerization of recycled monomers.

Molecular origin of the urethanase function toward *N*-aryl carbamates

To investigate the molecular origin of the urethanase function, we conducted a comparative structural analysis. The superimposed structures derived from GRASE and other screening algorithms revealed a conserved overall architecture, with root mean square deviation values consistently <1 Å, reflecting the limitation of global structure searches being unable to discern functional subtleties. Upon applying the GRASE framework, we suggest the variations in the local architecture surrounding the catalytic triad of the active enzymes, particularly within the lid domain (Fig. 4A). As shown in Fig. 4B, the nonfunctional proteins from global structure retrieving (e.g., F4 and F5) appeared to harbor narrow, constricted active sites that might impede the accommodation of bulky *N*-aryl carbamates such as TDA-DEG. By contrast, active proteins curated through GRASE were predicted to feature enlarged binding cavities that may facilitate van der Waals interaction stabilizations and enable more favorable orientations of larger aromatic groups. Those exceeding 1 U/mg in specific activity were predicted to have a highly open lid domain, forming a V-shaped architecture and potentially providing ample space for conducting substrate engagement and efficient catalysis. Certain lower-activity enzymes, although predicted to have enlarged active site pockets, appeared to feature relatively restricted lid domains. Despite the low activity of the enzymes identified through structural retrieval, their inclusion suggested considerable diversity in the architecture of the predicted lid domain. For example, GF5 and GF8, which displayed activities <0.1 U/mg, were predicted to have structurally compromised lid domains. GF5 may contain a single helix instead of the typical two-helix configuration in the lid domain, whereas GF8 was predicted to lack a lid domain. These structural variations suggest that although an open lid domain may be integral for facilitating substrate accommodation, the

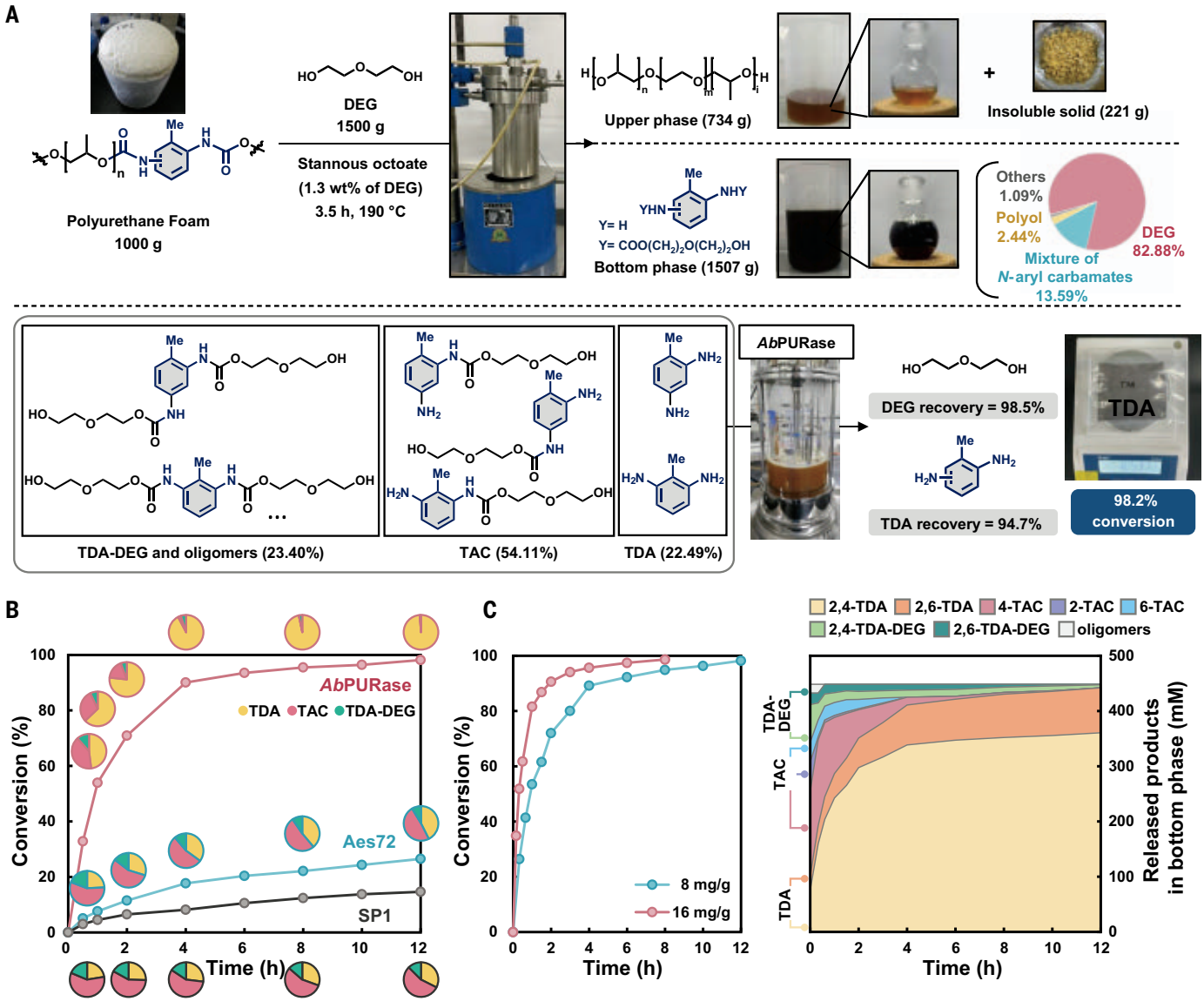
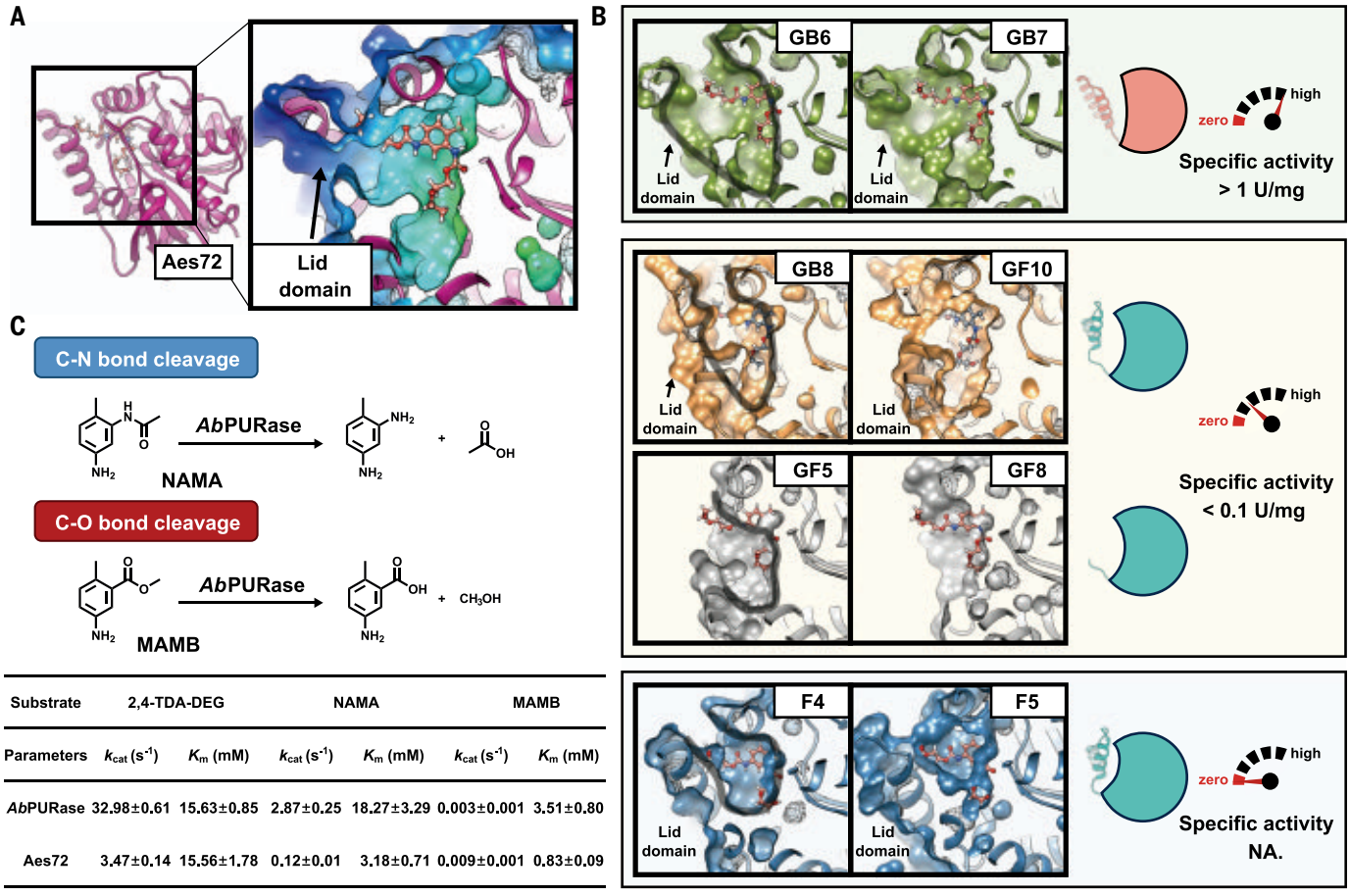


Fig. 3. Chemo-enzymatic depolymerization of BASF polyurethane foam. (**A**) Chemo-enzymatic depolymerization of one kilogram of polyurethane foam. Glycolysis yielded an upper phase enriched with EO-PO-EO copolyether polyols and a bottom phase composed of 8.6 M DEG and low-molecular-weight TDA dicarbamate. The insoluble fraction consisted mainly of additives. The pie chart shows the distribution of the compounds contained in the bottom phase. *Ab*PURase hydrolyzed the TDA dicarbamates in the bottom phase into TDA monomers with an enzyme loading of 8 mg/g. The TDA and DEG products were subsequently purified through rectification. (**B**) Enzymatic hydrolysis of the bottom phase derived from the glycolysis of 30 g of BASF polyurethane foam with 8 mg/g *Ab*PURase (50°C), Aes72 (37°C), and SP1 (30°C) in the bottom phase. Data are from one independent experiment. The pie charts illustrate the relative proportions of TDA-DEG, TAC, and TDA measured at each time point. (**C**) The left panel displays enzymatic hydrolysis of the bottom phase derived from the glycolysis of 1000 g of BASF polyurethane foam with 8 mg/g and 16 mg/g *Ab*PURase at 50°C. The right panel shows temporal variations of the TDA-DEG, TAC, and TDA compositions during enzymatic hydrolysis with 8 mg/g enzyme loading. Data are from one independent experiment.

absence of this critical architecture would severely compromise enzyme functionality.

The critical role of lid domains in modulating functions is commonly seen in the evolutionary adaptations of esterases and lipases, where catalysis in nonaqueous environments is facilitated by a structural transition of the lid domain from a closed to an open conformation (known as interfacial activation) (*34*). Unlike SP1, which belongs to the amidase signature family and uses a Ser-(*cis*)Ser-Lys catalytic triad, *Ab*PURase belongs to the α/β-hydrolase fold-3 (ABH3) superfamily and uses the hydrolase catalytic triad of Ser-His-Asp/Glu. Phylogenetic inference and sequence similarity network analysis suggested phylogenetic independence between these enzymes (fig. S12). Function annotation classified *Ab*PURase and other GRASE-grasped urethanases as carboxylic-ester hydrolases (EC 3.1.1). On the basis of these observations, we validated the esterase activity of *Ab*PURase by hydrolyzing methyl 5-amino-2-methylbenzoate (MAMB) (Fig. 4C). However, *Ab*PURase also exhibited amidase activity, as demonstrated by its cleavage of the C–N bond in *N*-(5-amino-2-methylphenyl)acetamide (NAMA), which was three orders of magnitude greater than its esterase activity (fig. S13). The kinetics analysis suggested that the high performance of *Ab*PURase toward urethane bonds may be attributed primarily to its strong amidase activity.

More crucially, this high activity was retained even in DEG solvents. Up to 6 M DEG, *Ab*PURase exhibited only marginal variation



Substrate	2,4-TDA-DEG		NAMA		MAMB	
Parameters	k_{cat} (s^{-1})	K_m (mM)	k_{cat} (s^{-1})	K_m (mM)	k_{cat} (s^{-1})	K_m (mM)
*Ab*PURase	32.98±0.61	15.63±0.85	2.87±0.25	18.27±3.29	0.003±0.001	3.51±0.80
Aes72	3.47±0.14	15.56±1.78	0.12±0.01	3.18±0.71	0.009±0.001	0.83±0.09

Fig. 4. Structural comparison among the enzymes and kinetic parameters. (**A**) Surface structure of Aes72 and its lid region. (**B**) Local structural differences among the active-site pockets of various enzymes highlighting the crucial role of the lid structure in catalytic performance. GF10 structure was obtained from the crystal structure (Protein Data Bank ID: 7NB5), and other enzyme structures were predicted using the AlphaFold3. (**C**) Comparison of Michaelis-Menten kinetic parameters for *Ab*PURase and Aes72 with the substrates 2,4-TDA-DEG, NAMA, and MAMB. All data are presented as mean ± SD ($n = 3$).

in apparent melting temperature (T_m), inferring its structural robustness. By contrast, Aes72 underwent substantial destabilization, with its T_m decreasing by 15°C at 6 M DEG (fig. S14). Consistent with the experimental results, Gaussian accelerated molecular dynamics simulations indicated structural integrity maintenance in *Ab*PURase at 6 M DEG, whereas Aes72 displayed a structural unfolding trend, particularly in the lid region (fig. S15). Detailed structural insights suggest that a tightly packed hydrophobic core in *Ab*PURase formed by residue M38 interacting with lid-domain residues (I9, M12, L13, L35, F197, W271, F277, and I279) could help to prevent DEG infiltration. Conversely, the shorter side chain of the corresponding residue L37 in Aes72 could potentially reduce these interactions, which might facilitate DEG penetration and disruption of the lid domain. $Loop_{\alpha1\text{-}\alpha2}$ in the *Ab*PURase lid domain may be stabilized by multiple proline residues (P24, P25, and P27), whereas Aes72, containing only P25 in $loop_{\alpha1\text{-}\alpha2}$, appears more flexible, potentially promoting partial unfolding of its lid structure. These molecular features provide preliminary insights into the structural stability of *Ab*PURase in DEG-rich environments, although a detailed understanding of the associated mechanism will require further in-depth research.

Discussion

The sequence and structural features underpinning polyurethane hydrolysis have been insufficiently characterized, both within and beyond the currently explored sequence space. The recent successes achieved by representation-based strategies, exemplified by approaches such as CLEAN and Pythia-Pocket, highlight the utility of learned latent embeddings for obtaining more reliable functional inference than conventional sequence- or structure-based comparisons because they capture high-dimensional biochemical features that are often overlooked by the traditional methods. However, subtle differences in catalytic properties may be obscured when embeddings are derived from global representations by using average pooling over the entire protein sequence. Pythia-Pocket addresses this by specifically embedding active-site residues rather than full sequences, enabling finer-grained functional discrimination even among distantly related homologs, as demonstrated by functionally active enzymes identified in this study that shared as little as 34% identity with the reference template. This capability is particularly beneficial for the identification of newly characterized enzymes, as exemplified by plastic-degrading enzymes, directly from a known template without the need for manual expert curation process, thereby accelerating the discovery in functionally underexplored regions of protein spaces. Recognizing that catalytic efficiency is more valuable when accompanied by adequate structural robustness, we integrated Pythia, a neural network that predicts folding stability from the sequence-structure likelihood into the unified GRASE framework, allowing a balance between enzymatic activity and robustness for industrial enzyme discovery.

Enzymes such as *Ab*PURase, which are traditionally classified as esterases, were found to be primarily urethanases in this study. This

observation aligns with the broader body of research on enzyme promiscuity and moonlighting, suggesting that many enzymes, rather than being strictly specific, have a range of activities that have often been obscured by conventional classification methods. Therefore, the full functional potential of enzymes may be significantly underappreciated in current annotations. To address this complexity, future studies may focus on efficiently capturing and amplifying the subtle signals that underlie these hidden capabilities. Such efforts will not only deepen our understanding of enzyme function repertoires, but will also contribute to the development of biocatalysts with broad applicability across various industrial challenges.

REFERENCES AND NOTES

1. V. Tournier *et al.*, *Nature* **580**, 216–219 (2020).
2. Plastics Europe, "Plastics – the fast facts 2024" (2024); https://plasticseurope.org/knowledge-hub/plastics-the-fast-facts-2024/.
3. G. Rossignolo, G. Malucelli, A. Lorenzetti, *Green Chem.* **26**, 1132–1152 (2024).
4. S. R. Nicholson, N. A. Rorrer, A. C. Carpenter, G. T. Beckham, *Joule* **5**, 673–686 (2021).
5. J. Liu *et al.*, *Biotechnol. Adv.* **48**, 107730 (2021).
6. A. Raczyńska, A. Góra, I. André, *Biotechnol. Adv.* **77**, 108439 (2024).
7. J. Liu *et al.*, *Appl. Environ. Microbiol.* **90**, e0147723 (2024).
8. F. Di Bisceglie, F. Quartinello, R. Vielnascher, G. M. Guebitz, A. Pellis, *Polymers* **14**, 411 (2022).
9. Z. Jiang *et al.*, *J. Hazard. Mater.* **472**, 134493 (2024).
10. A. Delavarde *et al.*, *Prog. Polym. Sci.* **151**, 101805 (2024).
11. Y. Branson *et al.*, *Angew. Chem. Int. Ed.* **62**, e202216220 (2023).
12. H.-W. Engels *et al.*, *Angew. Chem. Int. Ed.* **52**, 9422–9441 (2013).
13. A. Kemona, M. Piotrowska, *Polymers* **12**, 1752 (2020).
14. D. Simón, A. M. Borreguero, A. de Lucas, J. F. Rodríguez, *Polym. Degrad. Stabil.* **109**, 115–121 (2014).
15. T. Vanbergen *et al.*, *ChemSusChem* **13**, 3835–3843 (2020).
16. G. Kiss, G. Rusu, F. Peter, I. Tănase, G. Bandur, *Polymers* **12**, 1533 (2020).
17. J. del Amo *et al.*, *ACS Sustain. Chem.& Eng.* **12**, 17479–17487 (2024).
18. V. Stepankova *et al.*, *ACS Catal.* **3**, 2823–2836 (2013).
19. T. Bayer *et al.*, *Angew. Chem. Int. Ed.* **63**, e202404492 (2024).
20. S. Trott, R. Bauer, H.-J. Knackmuss, A. Stolz, *Microbiology* **147**, 1815–1824 (2001).
21. S. Trott, S. Bürger, C. Calaminus, A. Stolz, *Appl. Environ. Microbiol.* **68**, 3279–3286 (2002).
22. K. Masaki *et al.*, *J. Biosci. Bioeng.* **130**, 115–120 (2020).
23. Y. Jia, J. Zhou, G. Du, J. Chen, F. Fang, *Food Biosci.* **36**, 100666 (2020).
24. N. Bittner, H. Nefzger, G. Behnken, G. Jaeger, S. Behnken, L. Reisky, "Novel urethanases for the enzymatic decomposition of polyurethanes," EPO Patent EP3587570A1 (2020).
25. N. Azu, A. Chappelle, *Enc. Toxicol.* **9**, 151–156 (2024).
26. M. van Kempen *et al.*, *Nat. Biotechnol.* **42**, 243–246 (2024).
27. H. Seo *et al.*, *Science* **387**, eadp5637 (2025).
28. ByteDance AML AI4Science Team *et al.*, Protenix-advancing structure prediction through a comprehensive AlphaFold3 reproduction. bioRxiv 631937 [Preprint] (2025); https://doi.org/10.1101/2025.01.08.631967.
29. T. Yu *et al.*, *Science* **379**, 1358–1363 (2023).
30. G. Steinkellner *et al.*, *Nat. Commun.* **5**, 4150 (2014).
31. J. Sun, T. Zhu, Y. Cui, B. Wu, *Innovation (Camb.)* **6**, 100750 (2025).
32. A. W. Franz, S. Buchholz, R. W. Albach, R. Schmid, *Green Carbon* **2**, 33–44 (2024).
33. J. del Amo *et al.*, *Materials* **17**, 2844 (2024).
34. M. J. Liszka, M. E. Clark, E. Schneider, D. S. Clark, *Annu. Rev. Chem. Biomol. Eng.* **3**, 77–102 (2012).
35. J. Sun, Y. Cui, B. Wu, "GNN-based Recommendation of Active and Stable Enzymes (GRASE)," Zenodo (2025); https://doi.org/10.5281/zenodo.15549656.

ACKNOWLEDGMENTS

The authors acknowledge the support of the Shanghai BASF Polyurethane Co. Ltd. **Funding:** This work was supported by the National Key R&D Program of China (grant 2024YFC3405500), the Strategic Priority Research Program of the Chinese Academy of Sciences (CAS) (grant XDB0810000), the National Natural Science Foundation of China (grants 32225002, 32170033, and 32422001), Fundamental Research Funds for the Central Universities (grant buctrc202539), the Youth Innovation Promotion Association CAS (grant 2022086), the Postdoctoral Fellowship Program of CPSF (fellowship GZB20240808), and the Beijing Municipal Science & Technology Project, China (grant Z241100007724009). **Author contributions:** Y.Chen and K.S. performed biochemical experiments and collected the data. J.Y.S. developed the Pythia and Pythia-Pocket models. T.Z., X.X., C.D., R.L., and J.R. contributed to data preparation and statistical analyses. R.L., X.L., and W.-C.G. performed process modeling and operating cost analysis. W.S. performed phylogenetic inference and sequence similarity network analysis. Y.C. performed structural modeling and analysis of enzymes and drafted the manuscript with input from all authors. Y.C. and B.W. secured funding, interpreted results, and supervised the study. All authors revised the manuscript and accepted its final version. **Competing interests:** Y.Chen, Y.Cui, and B.W. have filed a patent application that covers the enzyme sequences in the paper. The remaining authors declare no competing interests. **Data and materials availability:** Code for Pythia-Pocket and Pythia are freely accessible at Zenodo (*35*) and GitHub (https://github.com/Wublab/Pythia/). All data used in this analysis are free to access and are available in the main text or the supplementary materials. A complete list of the nucleotide sequences used in this study is provided in data S1.

License information: Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adw4487
Materials and Methods; Supplementary Text S1 to S4; Figs. S1 to S36; Tables S1 to S14; References (*36*–*48*); MDAR Reproducibility Checklist; Data S1

Submitted 2 February 2025; accepted 21 August 2025

10.1126/science.adw4487

EPIDEMIOLOGY

Collateral effects of COVID-19 pandemic control on the US infectious disease landscape

Tobias S. Brett[1]* and Pejman Rohani[1,2,3]

Using data from the United States Centers for Disease Control and Prevention (CDC) disease surveillance systems, we sought to quantify the indirect effects of the COVID-19 pandemic, and the possibility of lack of exposure to common pathogens resulting in immune deficits. Clustering analysis on pandemic-era time-series data identified pathogen groupings according to transmission mechanism. Counterfactual analysis, using Bayesian structural time-series (BSTS) modeling, confirmed that infectious diseases that are directly transmitted via airborne droplets (aerosols) experienced the greatest disruption to transmission. By contrast, sexually transmitted infections (STIs) experienced a smaller transient disruption, and increasing trends in incidence prepandemic appear to have been curtailed. Using epidemiological theory, we demonstrate that the observed magnitudes and durations of notifications deficits were determined by fundamental disease system properties, namely, the serial interval, basic reproductive number, and susceptible recruitment.

The COVID-19 pandemic, caused by the emergence of the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) coronavirus (*1*), triggered an unprecedented global disruption to human society. Although governmental response measures and individual behavioral changes were targeted at curtailing the spread of SARS-CoV-2 (*2*), they also influenced co-circulating, sympatric infectious diseases (*3*–*10*). Of particular concern was the prospect that the implementation of social distancing measures would lead to gaps in population immunity against these diseases, with the potential for serious rebound in incidence for some infections after the relaxation of control measures (*11*–*15*). These fears, motivated by epidemiological theory (*16*, *17*), were borne out in part by the "triple-demic" of SARS-CoV-2, respiratory syncytial virus (RSV), and influenza virus (Fig. 1, A and B) during the winter of 2022–2023 (*18*) and may be contributing to a resurgence of pertussis in 2024 and 2025 (Fig. 1D) (*19*). Additional concerns were raised about potential disruption to disease surveillance efforts, routine vaccination programs (*20*), patient care–seeking behavior (*21*), and health care system capacity (*22*). Ecological competition by a novel pathogen may also have disrupted the community of human infectious diseases (*23*, *24*). The interpretation of post-COVID disease incidence trends is complicated by pandemic-spurred improvements in disease surveillance methodologies (*25*, *26*), spatial variation in disease incidence, and state government responses (Fig. 1, E to J).

The short-term effects of social distancing measures on sympatric infectious diseases have been examined for specific pathogens, including influenza (*11*, *27*, *28*), RSV (*11*, *29*), polio (*30*), tuberculosis (*31*, *32*), pertussis (*10*, *33*, *34*), and enterovirus D68 (*35*). Though valuable in providing epidemiological insights, most of these studies were performed at the height of the pandemic (2020–2021) with the aim of providing "scenario analyses" to project possible postpandemic implications for specific pathogens and diseases. Here, we present a retrospective analysis of surveillance data spanning the pandemic, with the aim of identifying patterns in the impacts of pandemic response measures on a suite of sympatric infectious diseases in the USA.

We carried out a comparative study using time-series data collected by the US Centers for Disease Control and Prevention and made publicly available through the National Notifiable Disease Surveillance System (NNDSS) and FluView (*36*, *37*). From the two sources, we collated a dataset of 24 infectious diseases amenable to statistical analysis.

Clustering analysis

To identify which groupings of diseases were similarly affected by the COVID-19 pandemic, we performed a clustering analysis on the national-level weekly notifications data from the start of 2019 onwards (Fig. 2; full time-series data are shown in fig. S1). We calculated the Spearman's ρ rank-correlation matrix, finding values of ρ that ranged from −0.47 (between syphilis and mumps) to 0.91 [between campylobacteriosis and Shiga toxin–producing *Escherichia coli* (STEC)] (Fig. 2A). We applied an agglomerative clustering algorithm to the correlation matrix, to construct a dendrogram of the hierarchical clustering of the data (Fig. 2B). On the basis of the silhouette score (fig. S2), we found that the data were most consistent with four clusters.

The clustering is suggestive of factors that might have helped shape the pandemic's effects on disease transmission (Fig. 2). The first cluster (shaded blue) comprises predominantly pathogens that are spread via airborne or aerosol transmission (with hepatitis B, spread by infected blood or bodily fluids, an outlier); one subgroup consists of vaccine-preventable childhood diseases (mumps, pertussis, invasive *H. influenzae*, and varicella) and the other, influenza A and B. The second (shaded pink) consists of the three sexually transmitted infections (STIs) in the dataset (gonorrhea, chlamydia, and syphilis), with coccidioidomycosis (environmentally transmitted) being an outlier. The third (shaded green) is more disparate but broadly consists of diseases that are not directly transmissible between humans, with one subgrouping consisting of bacteria spread fecal-orally (with malaria an outlier) and a second subgroup of tick-borne pathogens. The final group consists of two apparent outliers: hepatitis A, caused by a virus primarily spread primarily via fecal-oral transmission [which experienced a pronounced multistate outbreak from 2016 to 2019 (*38*)], and notifications of rabies in animals [the only nonhuman disease in the dataset, included for comparative purposes, which is subject to an ongoing vaccination campaign (*39*)]. In what follows, we will focus on the first three clusters, which we label according to predominant transmission mechanism: (i) airborne and aerosol-transmitted infections (ATIs), (ii) STIs, and (iii) environmentally transmitted infections and zoonotic spillover (ETI/Z). To verify that the clustering corresponds to effects of the pandemic and not prepandemic patterns, we repeated the clustering analysis using data pre-2020 (fig. S3). The resulting dendrogram was less reflective of transmission mechanism, although some features were preserved based on seasonal patterns of incidence (e.g., the STI cluster and a high correlation between influenza A and B).

Counterfactual analysis

To quantify the effects of the pandemic on sympatric disease notifications in different US states, we performed a counterfactual Bayesian structural time-series (BSTS) modeling analysis (*40*). The analysis was designed to estimate the discrepancy between observed notifications post-March 2020 and the counterfactual projected notifications predicted to have occurred had no pandemic taken place (see materials and methods). We fit the BSTS model to time-series data for each pairwise combination of US state and disease. Model validation was performed using posterior predictive checks, with 298 pairs passing the check (fig. S4).

[1]Odum School of Ecology, University of Georgia, Athens, GA, USA. [2]Department of Infectious Diseases, College of Veterinary Medicine, University of Georgia, Athens, GA, USA. [3]Center for Influenza Disease and Emergence Research (CIDER), Athens, GA, USA. ***Corresponding author. Email: tsbrett@uga.edu**

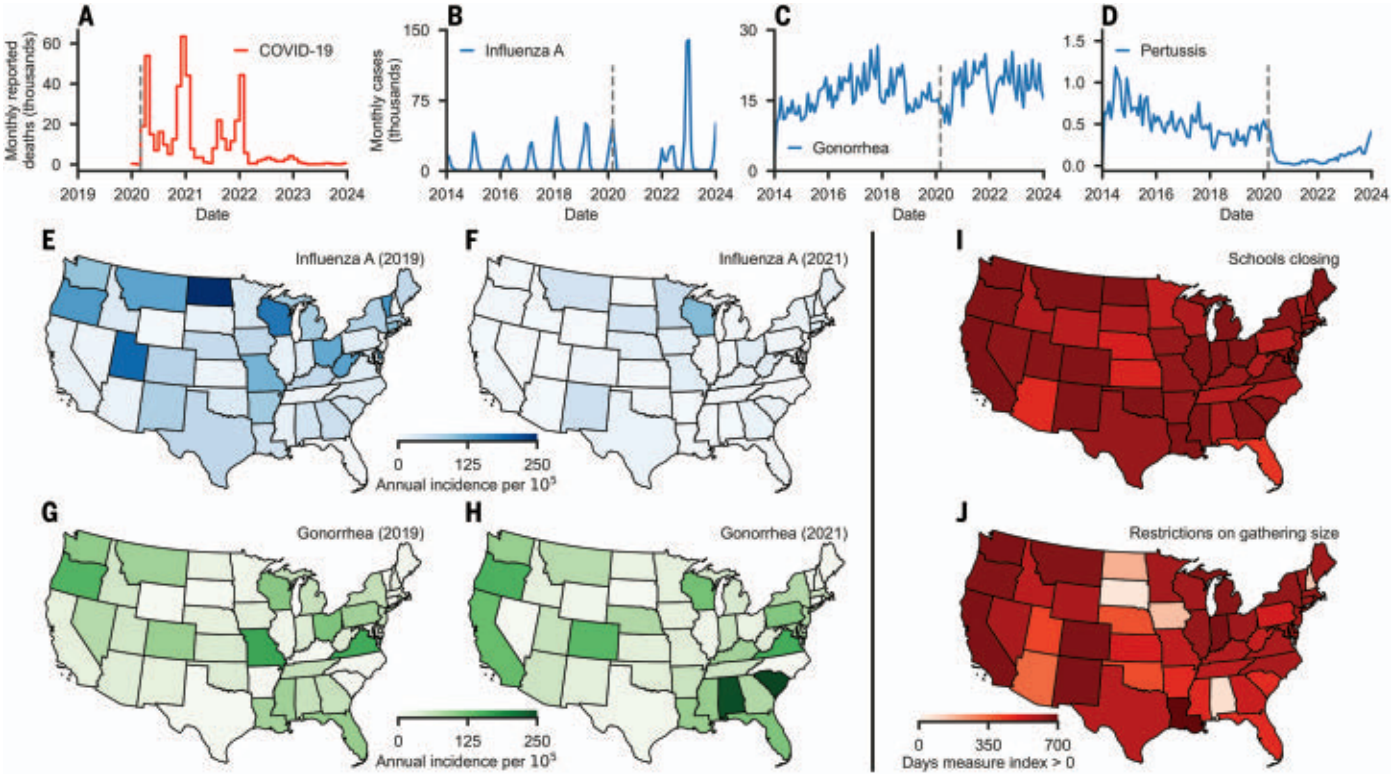


Fig. 1. Overview of pre– and post–COVID-19 notifications for select diseases. (**A**) Monthly recorded COVID-19 mortality. (**B** to **D**) Monthly notifications of cases of influenza A (B) gonorrhea (C), and pertussis (D). (**E** to **H**) On the state level, differences in notifications of influenza A were consistently lower in 2021 compared with 2019 (E and F). By contrast, changes in notifications of gonorrhea varied by state with some up (e.g., Alabama, California, South Carolina) and others down (e.g., Missouri, Nevada, Ohio). (**I** and **J**) The duration of government response measures also varied by state, as quantified by the Oxford COVID-19 Government Response Tracker (OxCGRT; see materials and methods). Although some level of disruption to education in schools remained across much of the country for almost 2 years (I), restrictions on gathering sizes were relaxed more quickly, and with greater spatial variation (J).

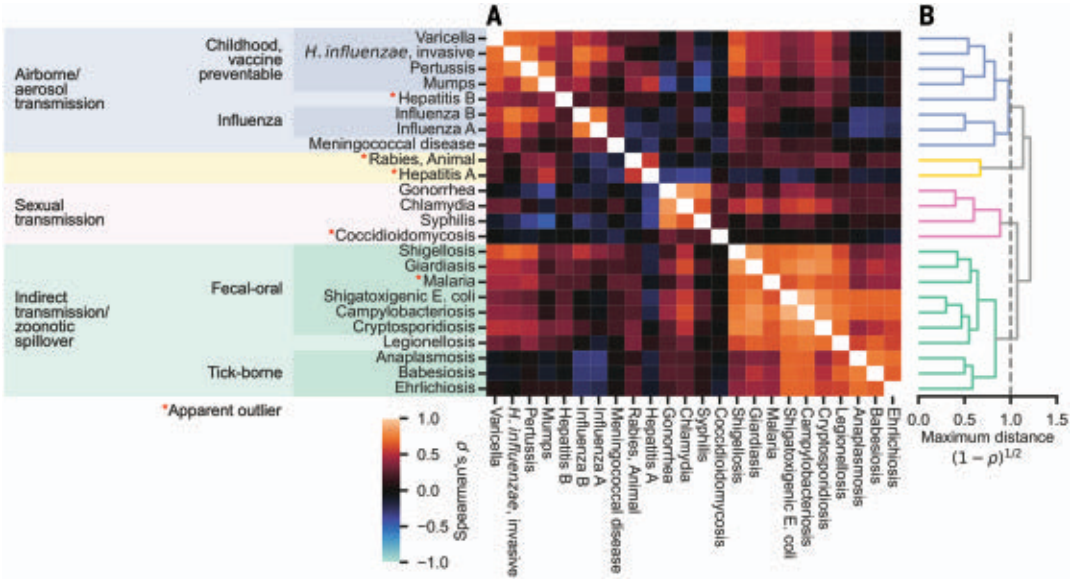


Fig. 2. Clustering analysis of sympatric disease time-series data from the US (2019–2024). (**A**) Spearman's rank correlation matrix for time-series data of weekly national case notifications for 24 infectious diseases (22 from the NNDSS plus lab-confirmed influenza A and B cases; see materials and methods). (**B**) Agglomerative clustering algorithm (*55*) applied to the correlation matrix, with the silhouette score (*56*) identifying four distinct clusters (fig. S2). Of these clusters, one (blue) primarily consists of airborne and aerosol-transmitted infections (ATIs). The second (pink) is of predominantly STIs. The third (green) is a broad cluster, which we label ETI/Z, that includes zoonotic tick-borne infections (anaplasmosis, babesiosis, and ehrlichiosis) and environmentally transmitted infections (ETIs). Diseases in this cluster do not depend on direct human contact for transmission. The final cluster (yellow) consists of two outliers, hepatitis A and rabies in animals.

The key measure used to interpret our results was the relative impact on disease notifications, ψ_t, given by $\psi_t = 100(y_t - \tilde{y}_t) / \tilde{y}_t$, which quantifies, as a percentage, the deviation of observed notifications y_t from counterfactual model projections, $\tilde{y}_t$ (Fig. 3, A and B). Similarly, the relative net impact, $\chi_t = 100 \sum_{s=0}^{t} (y_s - \tilde{y}_s) / \sum_{s=0}^{t} \tilde{y}_s$, quantifies the cumulative deviation of observed notifications from counterfactual model projections from the start of the pandemic up to time t. Related approaches have previously been used to quantify excess deaths attributable to the pandemic (*41*).

In Fig. 3, C to J, we present results for four diseases in New York state, representative of the three transmission clusters (Fig. 2A). Results for all other state-disease pairs are presented in figs. S5 to S10. In 2020, notifications of both influenza A (Fig. 3G) and pertussis (Fig. 3H) were substantially below counterfactual projections (posterior median annual relative annual impact −81.5% [−90.5%, −69.7%] and −76.6% [−81.8%, −71.2%], respectively), whereas the relative annual impact on shigellosis notifications (Fig. 3J) was −45.7% [−55.7%, −34.5%]; all ranges correspond to 95% highest density intervals (HDIs). Notifications of all three diseases subsequently rebounded with the annual relative impact 95% HDI containing 0% by the end of 2022. The return of both pertussis and shigellosis to projected levels has been gradual over the course of 2020–2022 (although there were explosive outbreaks of pertussis in New York during 2024, among other locations). By contrast, influenza A was largely absent until a sudden rebound in 2022 (Fig. 3G). Unlike the other three diseases shown, there was no apparent effect on weekly gonorrhea notifications in 2020 (Fig. 3I), with a relative annual impact of −1.2% [−8.5%, 6.5%], although notification reports were absent from early July until late November 2020. Subsequently, there was a small negative impact in 2022 and 2023, as notifications undershot projections that sustained an increasing pre-2020 trend. For all four diseases shown (Fig. 3, G to J), the relative net impact remains below zero at the end of 2023, corresponding to a net deficit in notifications.

Given the large number of state-disease pairs, we summarized our results by aggregating estimates of the relative net impact on each pathogen for each cluster (Fig. 4). We found that the median for all three clusters (with influenza excluded from the ATI cluster) remained below zero at the end of 2023 (Fig. 4A). Although following a similar trajectory to other ATIs, both influenza A and B were most influenced by the pandemic during 2021 (Fig. 4B). Subsequently, during 2022 and 2023, the influenza A rebound was positive in some states (in particular South Dakota, Hawaii, and Nebraska), whereas in others, the influenza seasons were smaller than expected (e.g., Michigan and Louisiana; fig. S11). By contrast, at the start of 2024, influenza B was suppressed in all states apart from Mississippi and South Carolina (fig. S12).

Broadly speaking, the effect of the pandemic and related response measures on other pathogens appears to be more strongly shaped by transmission mechanism of the pathogen concerned than state-level effects (Fig. 4, C to E). To understand whether state-specific differences in COVID-19 response measures could explain the interstate variation in the relative net impact on other pathogens, we formulated a multilevel regression model that we fit to predict the median relative net impact at the end of 2021, 2022, and 2023 using a dataset of government response measures as predictors (figs. S13 to S16). Except for the school closure coefficient for the ATI cluster, the posterior 95% HDI for each coefficient in 2021 included zero (fig. S14B). This indicates that our model was unable to use those predictors to explain interstate variation in relative net impact (χ_t). Our results do not imply that those response measures made no contribution to χ_t, given that the data were standardized and that we estimated nonzero offsets for

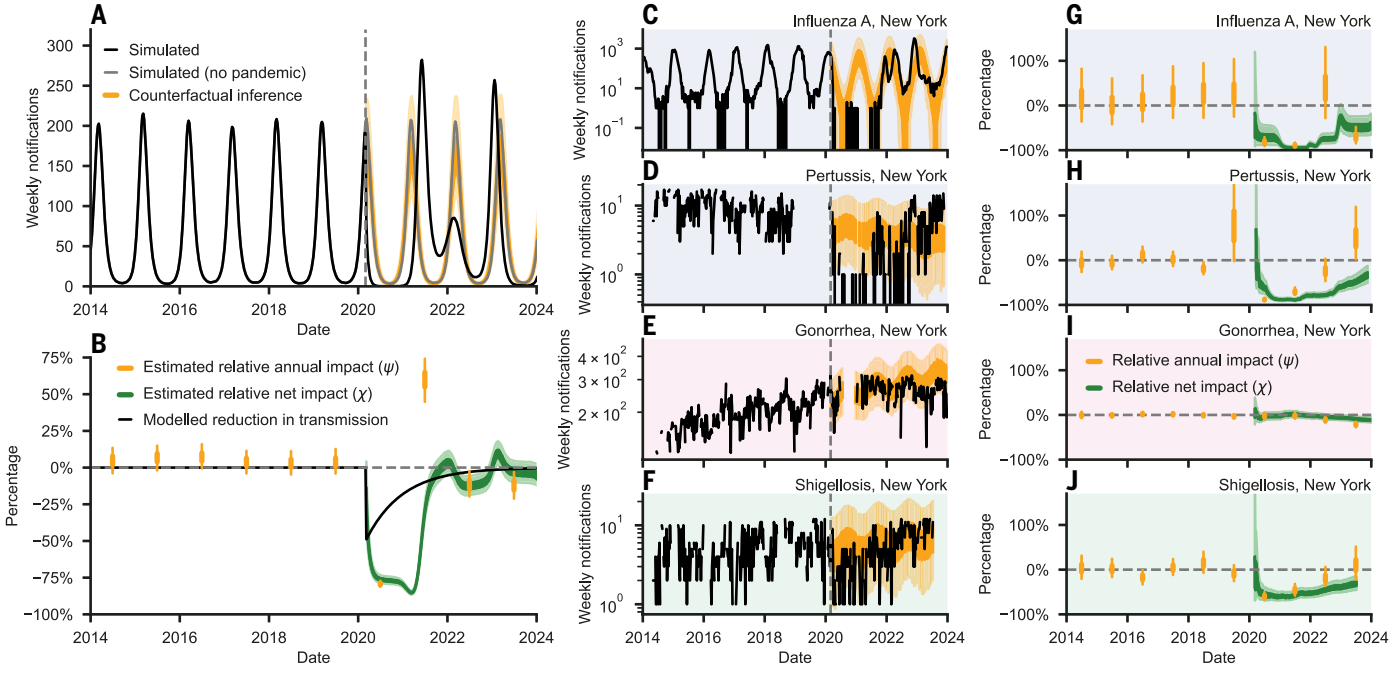


Fig. 3. Results of the counterfactual inference analysis for four diseases in New York. (**A**) Demonstration of counterfactual approach using simulated data. The model was fit to data pre-March 2020, with counterfactual projections (which assumed no pandemic effects) made for subsequent weeks. The counterfactual projections agree with simulated data that assumed no pandemic (gray line) but deviated from the simulated trajectory that assumed an initial 50% reduction in transmission. (**B**) The discrepancy between data post-2020 and the counterfactual projections allow us to quantify the relative impact, which lags the modeled reduction in transmission. Orange intervals show the relative annual impact for a given calendar year; green ribbons show the cumulative relative net impact since the pandemic start (dark and light bars/ribbons correspond to the 50% and 95% HDI, respectively). (**C** to **F**) Weekly observed notifications of influenza A, pertussis, gonorrhea, and shigellosis (black) are compared with counterfactual predictions from the BSTS model. (**G** to **J**) Relative annual and net impact of the pandemic. The annual impact calculations for 2020 only include data post-March.

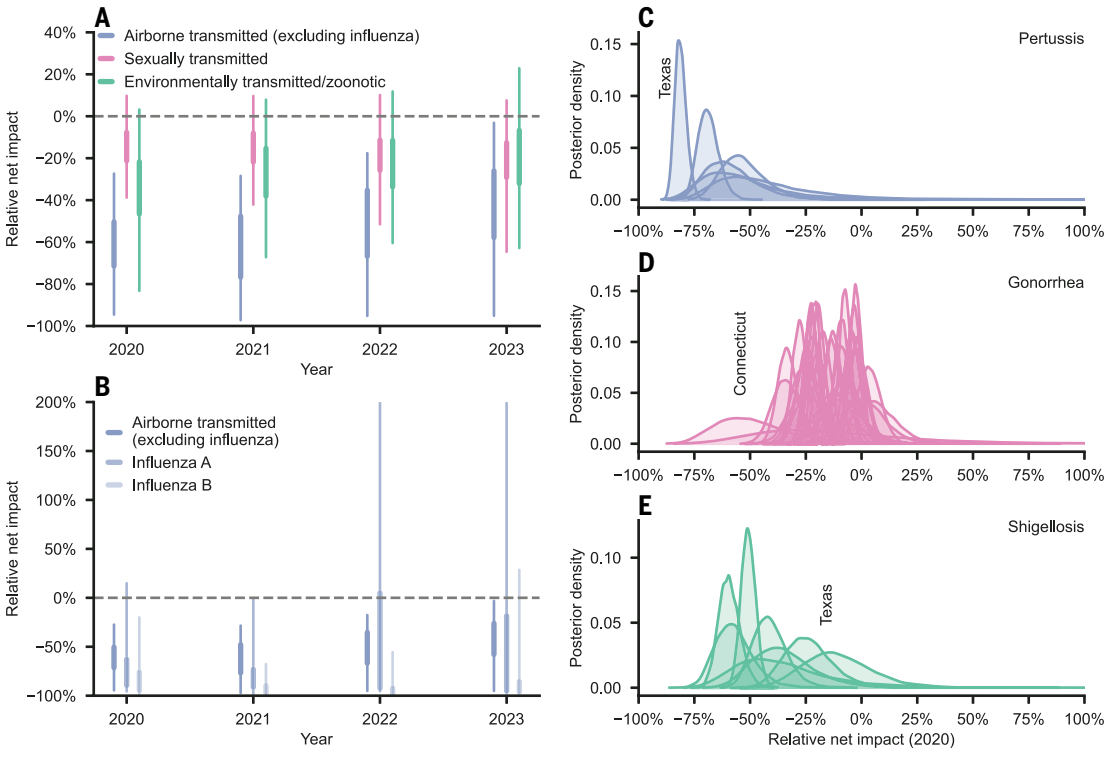


Fig. 4. Relative net impact by year and cluster. (**A**) Relative net impact at the end of each year for each transmission cluster. Densities computed by aggregating posterior samples for all state-disease pairs in a cluster (points indicate the median; bars correspond to 50% and 95% HDIs). Owing to the atypical relative net impact on influenza, we excluded it from the ATI cluster. (**B**) Comparison of the relative net impact on influenza with the rest of the ATI cluster diseases, using the same approach as for (A). For influenza A, the 95% HDI extends to 424% and 435% in 2022 and 2023, respectively. (**C** to **E**) State-level variation in relative net impact for pertussis (C), gonorrhea (D), and shigellosis (E).

each cluster (fig. S14C). Qualitatively similar results were observed for both 2022 and 2023 (figs. S15 and S16).

To explore the possible mechanisms underlying the differences between the two directly transmitted clusters (ATIs and STIs), we formulated and analyzed a simple pathogen transmission model (see materials and methods for details). We modeled the pandemic as causing a reduction in the transmission rate, which followed a decaying exponential, $c(t) = c_0\left(1-e^{-t/\tau}\right)$, where c_0 is the reduction at the start of the pandemic and τ determines the rate of return to the prepandemic transmission level (Fig. 5A). Following the methodology of our counterfactual analysis, we calculated the relative net impact and the return time, defined as the mean time for the relative net impact to return within 10% of zero. Using plausible parameter values, we simulated three diseases: pertussis (Fig. 5, B and C), influenza A (Fig. 5, D and E), and gonorrhea (figs. S17 and S18). We also simulated a set of parameter scenarios (Fig. 5H and figs. S19 and S20). For pertussis, we see a return time of about 4 years, followed by decaying oscillations in the net impact. For influenza, we see greater variation between stochastic replicates, with the net impact returning to zero after about 1 to 2 years. Basic epidemiological theory can explain the observed dynamics. Early in the pandemic, prevalence drops exponentially, $\frac{dI(t)}{dt} \approx -c_0\gamma I(t)$, depending on the severity of control measures, c_0, and the pathogen infectious period, $1/\gamma$. Later, as the transmission rate approaches prepandemic levels, the rebound is shaped by the basic reproductive number, R_0, and the rate of influx into the susceptible pool, κ, approximately following $\frac{dI(t)}{dt} \approx \gamma\kappa R_0 I(t)$. The parameter κ comprises the rate of unvaccinated births and loss of infection- and vaccine-derived immunity. Seasonal infectious diseases with low basic reproductive number, short infectious period, and short-lived immunity, such as influenza, are expected to undergo a sharp decline followed by a volatile rebound before returning to zero. These patterns are broadly in line with observations for influenza A (fig. S11). The return time is sensitive to nonlinearities in the transmission process, analogous to period doubling observed for vaccine-controlled seasonal diseases (*42*). For gonorrhea, we find that the observed relative net impact (for instance in New York; Fig. 3I) is most consistent with a smaller effect of the pandemic compared to the ATIs (both in the magnitude and duration of reductions in transmission rates; fig. S17). Furthermore, to match trends in the data, our modeling required that an increasing trend in transmission rates pre-2020 ended at the start of the COVID-19 pandemic.

Discussion

A limitation of counterfactual analyses is the assumption that data trends can be extrapolated from the training period (the prepandemic period in our study) into the counterfactual projection period (i.e., after March 2020). Therefore, a major caveat to the findings of our BSTS analysis is that it cannot, in isolation, distinguish between effects of the pandemic and other coincidental secular trends in transmission dynamics and surveillance after 2020. For instance, contemporary patterns of incidence may be affected by factors unrelated to the pandemic period, including changes in diagnostic methodology [e.g., increased usage of multiplex assays (*25*, *26*)], behavior [e.g., vaccine hesitancy (*43*)], and pathogen transmissibility and virulence (*44*). Another limitation is our choice of model structure. Given the large number of diseases studied, we opted to use a generalized linear model; however, appropriate disease-specific mechanistic transmission models could provide better interpretations of the data. Furthermore, as we were fitting to 6 years of data pre-2020, diseases with long interepidemic periods [such as mumps (*45*)] were difficult to fit with

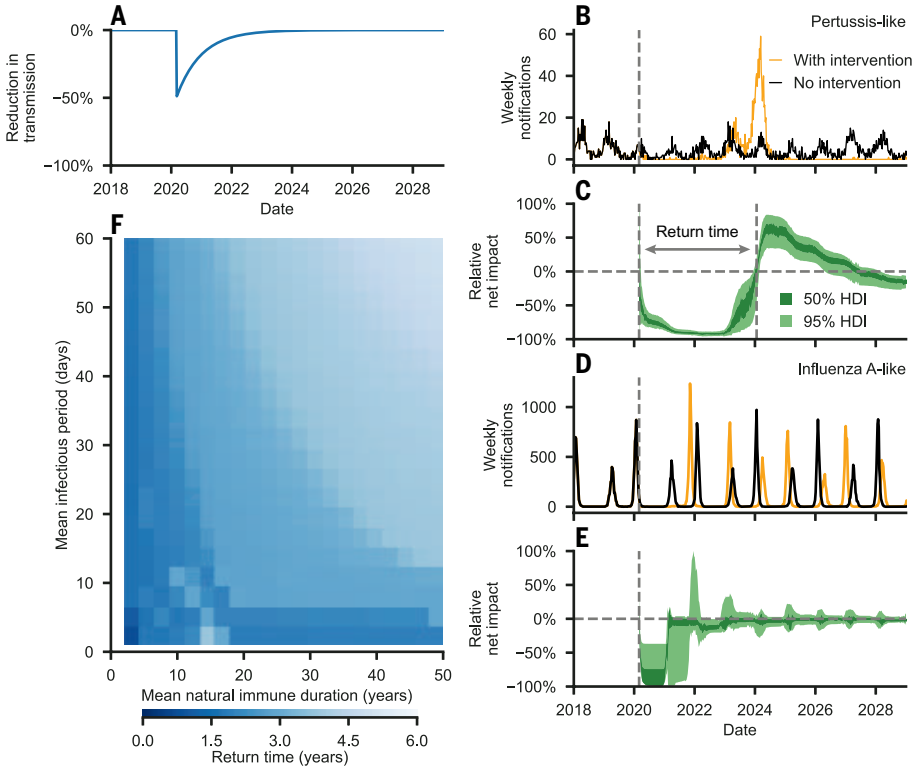


Fig. 5. Scenario analysis. (**A** to **E**) Simulated results using a stochastic SEIRV model (see materials and methods) subject to an initial 50% reduction in transmission rates (A). Results shown using plausible parameterizations for two ATIs, pertussis (B and C) and influenza A (D and E). (**F**) Dependence of the return time on the natural immune duration and infectious period for a pathogen with $R_0 = 4$. The return time was defined as the duration between the start of control measures and the mean relative net impact crossing −10% (C).

our nonmechanistic model. For instance, mumps and hepatitis A experienced pronounced outbreaks in several states during the prepandemic period (2014–2020), which led to rejection under statistical inclusion criteria (see materials and methods).

Our counterfactual analysis shows that notifications of some of the infectious diseases being investigated, in particular STIs, were largely unaffected by the pandemic (Fig. 4). From this observation, we infer that there was no generalized collapse in disease surveillance during the acute phase of the pandemic, and decreases in notifications of other diseases likely reflect actual reductions in transmission. Nevertheless, we cannot rule out the possibility of disproportionate disruption to the surveillance of specific diseases or in specific states.

The limited effect on notifications of the three STIs during 2020 and 2021 is puzzling, given the severe lockdown measures imposed and their effect on SARS-CoV-2 and ATI transmission. Our analysis of a gonorrhea-like transmission model suggests that even a moderate (initially 25%) short-lived reduction in transmission rates should have had a greater discernible net impact than observed for most state-disease pairs in the STI cluster (fig. S17). The longer serial interval of STIs relative to ATIs alone cannot explain the limited drop in notifications during 2020 (figs. S17 to S20). Another unexplained observation in these data (2014–2024) is that increasing trends in notifications that predated the pandemic have been absent since 2020, a phenomenon repeated across states and STIs in the study. Our results point to a persistent change in STI transmission dynamics that appears to us unexplained.

It is not surprising that ATIs were most affected by the pandemic, given the types of recommended control measures (e.g., mask wearing, reductions in mass gatherings, school and workplace closures) that were implemented to stop the spread of an aerosol-transmitted virus such as SARS-CoV-2. However, the delay in the rebound of these infectious diseases is noteworthy. For instance, in the US, despite public health interventions, contact rates were still sufficient for the rapid spread of the alpha, delta, and omicron variants in 2021 and influenza A in winter 2022 (*46*). By contrast, over the same period, the incidence of influenza B, invasive *H. influenzae*, mumps, pertussis, and varicella all remained below expected levels. Epidemiological modeling indicated a number of possible mechanisms (Fig. 5 and fig. S19). Specifically, the suppressed infectious diseases have longer serial intervals (resulting in slower epidemiological growth rates), lower transmissibility, and longer durations of immunity (and therefore slower susceptible recruitment). Regardless of parameterization, our modeling suggests that the effect of the pandemic is unlikely to extend beyond about 6 years. Our results point to a need for more systematic study of the underlying mechanisms. Indeed, the disruptive effects of the pandemic on the circulation of other infectious diseases provides a valuable opportunity to better understand the epidemiology of these pathogens.

Of the ATIs, it is not clear why influenza A and B notifications were disproportionately affected (Fig. 4B). Perhaps this is a result of disrupted international travel, which is known to play an important role in the circulation and seeding of novel influenza antigenic variants (*47*, *48*); however, curiously, influenza A was less affected by this perturbation than influenza B, despite the higher frequency of variant introductions before the pandemic (*48*). Alternatively, viral interference after SARS-CoV-2 infection and vaccination may have affected sympatric viruses (*49*). Apart from SARS CoV-2, influenza viruses are the only viruses in our study that infect the respiratory tract; therefore, it is possible that they experienced the greatest interference from the emergence of SARS-CoV-2.

Our quantification of the pandemic's impact has focused on the USA, where SARS-CoV-2 has circulated continuously since early 2020. However, the pandemic left a qualitatively similar footprint on notifications data for influenza in Australia (fig. S21) and pertussis in New Zealand (fig. S22), countries that adopted a zero-COVID strategy and largely suppressed SARS-CoV-2 circulation until late 2021. Timing of interventions relative to epidemic phase was found to affect the return time, with simulations where transmission reductions coincided with the start of the epidemic season demonstrating longer disruption (fig. S20).

Our findings are consistent with available studies of other populations [US military personnel (*50*)] and disease systems [*Salmonella* in Germany (*51*)]. An intriguing study of seasonal endemic respiratory viruses in Washington state, including a number of viruses not part of the NNDSS, found that the non-enveloped viruses in the study (human rhinoviruses, enteroviruses, and adenoviruses) were the first to rebound after the lifting of lockdown measures (*52*). The authors speculate that virological factors, such as environmental stability and ability to penetrate face masks, may have facilitated their rapid spread. Alternatively, our results suggest that short serial intervals and short-lasting immunity ($\lesssim$ 2 years) are sufficient to explain their rapid rebound.

In 2025, notifications of the ATIs largely rebounded to prepandemic levels as expected (*11*, *13*). It was unclear whether the rebound in these

diseases would exceed the incidence prevented by pandemic control measures, owing to the immunity gap created by reduced pathogen circulation and missed routine vaccinations. Our results show that cumulative notifications of most ATIs in most states since 2020 remained at a net deficit compared with counterfactual projections at the start of 2024. Given that we do not find evidence for a substantial collapse in disease reporting during 2020, we expect this deficit to reflect an overall reduction in the burden of disease. Although governmental-response measures to control the pandemic undoubtedly had adverse psychological, economic, and societal consequences (*53*), considered solely through the lens of infectious disease notifications, lockdown measures were beneficial for human health and did not lead to a collateral excess of sympatric disease. Extrapolating from disease notifications to the total number of infections in the population, and the implications for population immunity, can only be done on a disease-by-disease basis and remains a pressing need for further study such as through serological surveillance (*54*) and pathogen transmission modeling (*45*).

REFERENCES AND NOTES

1. N. Zhu *et al.*, *N. Engl. J. Med.* **382**, 727–733 (2020).
2. T. Hale *et al.*, *Nat. Hum. Behav.* **5**, 529–538 (2021).
3. L. Rodgers *et al.*, *Clin. Infect. Dis.* **73** (Suppl 1), S110–S117 (2021).
4. K. Hirae, T. Hoshina, H. Koga, *Int. J. Infect. Dis.* **128**, 265–271 (2023).
5. J. Hatoun, E. T. Correa, S. M. A. Donahue, L. Vernacchio, *Pediatrics* **146**, e2020006460 (2020).
6. A. Ullrich *et al.*, *Lancet Reg. Health Eur.* **6**, 100103 (2021).
7. J. A. McBride, J. Eickhoff, E. R. Wald, *Pediatr. Infect. Dis. J.* **39**, e449–e452 (2020).
8. M. J. Geng *et al.*, *Nat. Commun.* **12**, 6923 (2021).
9. S. Matczak *et al.*, *Euro Surveill.* **27**, 2100933 (2022).
10. E. Tessier *et al.*, *BMC Public Health* **22**, 405 (2022).
11. R. E. Baker *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30547–30553 (2020).
12. M. W. Fong, N. H. L. Leung, B. J. Cowling, P. Wu, *Emerg. Infect. Dis.* **27**, 1525–1527 (2021).
13. K. Messacar *et al.*, *Lancet* **400**, 1663–1665 (2022).
14. A. G. Feldman, S. T. O'Leary, L. Danziger-Isakov, *Clin. Infect. Dis.* **73**, 1920–1923 (2021).
15. D. N. Durrheim *et al.*, *Nat. Med.* **27**, 360–361 (2021).
16. A. Handel, I. M. Longini Jr., R. Antia, *Proc. Biol. Sci.* **274**, 833–837 (2007).
17. T. D. Hollingsworth, D. Klinkenberg, H. Heesterbeek, R. M. Anderson, *PLOS Comput. Biol.* **7**, e1001076 (2011).
18. E. Rios-Guzman *et al.*, *Nat. Commun.* **15**, 3374 (2024).
19. C. Rodrigues *et al.*, *Euro Surveill.* **29**, 2400459 (2024).
20. K. Causey *et al.*, *Lancet* **398**, 522–534 (2021).
21. R. Moynihan *et al.*, *BMJ Open* **11**, e045343 (2021).
22. C. Arsenault *et al.*, *Nat. Med.* **28**, 1314–1324 (2022).
23. P. Rohani, C. J. Green, N. B. Mantilla-Beniers, B. T. Grenfell, *Nature* **422**, 885–888 (2003).
24. D. A. Vasco, H. J. Wearing, P. Rohani, *J. Theor. Biol.* **245**, 9–25 (2007).
25. J. A. Hay *et al.*, *Science* **373**, eabh0635 (2021).
26. R. E. Baker *et al.*, *Nat. Rev. Microbiol.* **20**, 193–205 (2022).
27. C. M. Zipfel, V. Colizza, S. Bansal, *Vaccine* **39**, 3645–3648 (2021).
28. S. S. Lee, C. Viboud, E. Petersen, *Int. J. Infect. Dis.* **122**, 1002–1004 (2022).
29. Z. Zheng, V. E. Pitzer, E. D. Shapiro, L. J. Bont, D. M. Weinberger, *JAMA Netw. Open* **4**, e2141779–e2141779 (2021).
30. D. A. Kalkowska *et al.*, *Vaccine* **41** (Suppl 1), A12–A18 (2023).
31. C. F. McQuaid *et al.*, *Eur. Respir. J.* **56**, 2001718 (2020).
32. L. Cilloni *et al.*, *EClinicalMedicine* **28**, 100603 (2020).
33. M. Briga, E. Goult, T. S. Brett, P. Rohani, M. Domenech de Cellès, *Nat. Commun.* **15**, 921 (2024).
34. H. He *et al.*, *Vaccine* **40**, 6956–6962 (2022).
35. S. W. Park *et al.*, *Epidemics* **46**, 100736 (2024).
36. US Centers for Disease Control and Prevention (CDC), NNDSS Data Tables [Internet]. 2023. Available from: https://data.cdc.gov/browse?q=NNDSS
37. Centers for Disease Control & Prevention, Weekly U.S. Influenza Surveillance Report 2024; https://www.cdc.gov/fluview/index.html [cited 11 February 2024].
38. M. A. Foster *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **68**, 413–415 (2019).
39. X. Ma *et al.*, *J. Am. Vet. Med. Assoc.* **260**, 1157–1165 (2022).
40. K. H. Brodersen, F. Gallusser, J. Koehler, N. Remy, S. L. Scott, *Ann. Appl. Stat.* **9**, 247 (2015).
41. D. M. Weinberger *et al.*, *JAMA Intern. Med.* **180**, 1336–1344 (2020).
42. D. J. Earn, P. Rohani, B. M. Bolker, B. T. Grenfell, *Science* **287**, 667–670 (2000).
43. D. A. Salmon, M. Z. Dudley, J. M. Glanz, S. B. Omer, *Vaccine* **33** (suppl. 4), D66–D71 (2015).
44. A. Sasaki, S. Lion, M. Boots, *Nat. Ecol. Evol.* **6**, 51–62 (2022).
45. D. V. Gokhale, T. S. Brett, B. He, A. A. King, P. Rohani, *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2207595120 (2023).
46. T. S. Brett, P. Rohani, *PNAS Nexus* **1**, pgac159 (2022).
47. T. Bedford *et al.*, *Nature* **523**, 217–220 (2015).
48. Z. Chen *et al.*, *Science* **386**, eadq3003 (2024).
49. J. Piret, G. Boivin, *Emerg. Infect. Dis.* **28**, 273–281 (2022).
50. B. Manuelpillai, B. Lopman, C. R. Doran, C. Porter, *J. Infect. Dis.* **231**, e1075–e1079 (2025).
51. K. Nelson, B. Lopman, *Science* **377**, 33–34 (2022).
52. A. C. Perofsky *et al.*, *Nat. Commun.* **15**, 4164 (2024).
53. H. G. Nijhuis, L. J. G. van der Maesen, *Int J Soc Qual.* **11**, v–xx (2021).
54. C. J. E. Metcalf *et al.*, *Lancet* **388**, 728–730 (2016).
55. T. Hastie, R. Tibshirani, J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, 2009), vol. 2.
56. P. J. Rousseeuw, *J. Comput. Appl. Math.* **20**, 53–65 (1987).
57. T. Brett, Data for: Collateral effects of COVID-19 pandemic control on the US, Zenodo (2025); https://doi.org/10.5281/zenodo.16772531.

ACKNOWLEDGMENTS

We thank B. Grenfell and the Rohani lab for comments on this work. **Funding:** National Institutes of Health (R21AI171509 (P.R., T.B.). National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under contract no. 75N93021C00018 (NIAID Centers of Excellence for Influenza Research and Response, CEIRR) (P.R.). **Author contributions:** Conceptualization: T.B., P.R. Methodology: T.B., P.R. Investigation: T.B. Visualization: T.B. Funding acquisition: P.R. Project administration: P.R. Supervision: P.R. Writing – original draft: T.B. Writing – review & editing: T.B., P.R. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data and code to reproduce results have been deposited in Zenodo (*57*) (https://doi.org/10.5281/zenodo.16772531). **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adw4964
Materials and Methods; Figs. S1 to S22; Tables S1 and S2; References (*58*–*71*);
MDAR Reproducibility Checklist

Submitted 5 February 2025; accepted 8 August 2025

10.1126/science.adw4964

CATALYSIS

Trace-level halogen blocks CO_2 emission in Fischer-Tropsch synthesis for olefins production

Yi Cai[1,2]†, Maolin Wang[3]†, Shu Zhao[4]†, Xi Liu[5]†, Junzhong Xie[3]†, Xing-Wu Liu[1,3,6]*, Yao Xu[3], Jie Zhang[3], Lingzhen Zeng[3], Fei Qian[1,2], Zirui Gao[3], Zeyan Cen[3], Xingchen Liu[1], Hong Wang[6], Bingjun Xu[3], Graham J. Hutchings[7], Yong Yang[1,6], Yong-Wang Li[6], Xiao-Dong Wen[1,2]*, Ding Ma[3]*

Sustainable production of fuels and olefins from syngas (carbon monoxide and hydrogen) through the Fischer-Tropsch synthesis process requires catalysts that deliver high selectivity, industrial productivity, and minimal carbon dioxide (CO_2) emissions. Current industrial iron catalysts form substantial CO_2 by-product that limits carbon efficiency. We report that introducing trace amounts [parts per million (ppm) level] of halogen-containing compounds into the feed gas can suppress CO_2 formation using iron-based catalysts and boost olefin selectivity over paraffin and olefin productivity. Cofeeding 20 ppm bromomethane over an iron carbide catalyst decreased CO_2 selectivity to <1% and increased olefin selectivity to ~85% among all carbon-containing products. Surface-bound halogens modulated the catalyst surface structure and selectively inhibited pathways responsible for CO_2 generation and olefin hydrogenation. This strategy provides a simple, scalable, and broadly applicable route for carbon-efficient syngas conversion.

Developing efficient catalytic systems for direct α-olefin production from syngas (CO and H_2) through Fischer-Tropsch synthesis (FTS) is essential to address the growing global demand for olefins and is of importance for regions with limited petroleum resources (*1*–*3*). A key challenge is the design of low-cost, non–noble-metal catalysts that operate under moderate conditions that also minimize CO_2 emissions to ensure sustainable carbon utilization (*4*–*8*). Among current approaches, iron-based Fischer-Tropsch to olefins (FTO) catalysts (*9*–*11*) and oxide-zeolite (OX-ZEO) tandem systems (*2*, *12*, *13*) have emerged as leading candidates. However, iron-based catalysts often yield only moderate olefin selectivity, whereas OX-ZEO systems suffer from high CO_2 selectivity (32 to 45%) and limited olefin space-time yields (*12*). These limitations constrain overall carbon efficiency and hinder their large-scale industrial implementation. Thus, there is an urgent need for new catalytic systems that combine ultralow CO_2 selectivity, high α-olefin space-time yields, and robust long-term stability.

Iron-based catalysts are particularly attractive for FTS because of their low cost, natural abundance, tunable product distribution, and high space-time yields of the products. Currently, more than two-thirds of the global FTS capacity (15.70 million tons per year) relies on iron catalysts. However, one drawback is their intrinsic promotion of the water-gas shift (WGS) and Boudouard reactions, resulting in excessive CO_2 formation (typically showing a selectivity of 18 to 35%) that reduces carbon utilization efficiency (*14*, *15*). The formation of CO_2 not only contributes to greenhouse gas emissions but also depletes valuable carbon feedstocks.

Suppressing CO_2 formation in Fe-based FTS remains a major challenge. Previous strategies to inhibit the WGS reaction have focused on limiting water readsorption by shielding iron catalysts with hydrophobic silica coatings (*11*) or graphene layers (*8*), which effectively reduces CO_2 selectivity to <13%. More recently, modified pure-phase iron carbide catalysts have achieved CO_2 selectivity as low as ~10% (*10*, *16*), but only under conditions of low CO conversion and limited olefin productivity (~0.6 $g \cdot g_{cat}^{-1} \cdot hour^{-1}$) (*10*). Therefore, the development of catalytic systems that concurrently minimize CO_2 formation and maximize α-olefin productivity under practical conditions remains a critical and unmet challenge.

We demonstrate that the introduction of parts per million (ppm) concentrations of halomethanes into the syngas feed fundamentally transforms the catalytic behavior of Fe-based FTS systems. Specifically, cofeeding 20 ppm of bromomethane (CH_3Br) effectively suppressed CO_2 formation, reducing CO_2 selectivity to <1% and drastically enhancing the olefin/paraffin (o/p) ratio from 1.3 (on unpromoted χ-Fe_5C_2) to as high as 13. Under optimized conditions, this strategy achieved an olefin space-time yield of 1.08 $g \cdot g_{cat}^{-1} \cdot hour^{-1}$ with an unprecedented olefin selectivity of 85.2%, calculated relative to all carbon-containing products. This simple and easily implementable halomethane-promotion strategy is compatible with a broad range of Fe-based FTS catalysts, including commercial formulations. Mechanistic insights from transient kinetic analyses and density functional theory (DFT) calculations reveal that surface-bound halogens modulate the properties of the catalyst and selectively inhibit key side reactions—specifically, the Boudouard reaction and WGS reaction—thereby suppressing CO_2 generation. We introduce a simple yet robust and broadly applicable strategy to achieve sustainable and highly efficient olefin or liquid fuels production with exceptional carbon efficiency.

The impact of halogen on the Fe-based FTS process

A series of iron carbide catalysts (ε-Fe_2C, θ-Fe_3C, *h*-Fe_7C_3, and χ-Fe_5C_2) were synthesized following established protocols (*17*, *18*). The unpromoted χ-Fe_5C_2 catalyst, one of the most active phases for FTS, exhibited a CO_2 selectivity of 31.4% at a CO conversion of 93.3% (Fig. 1A) (*19*). The primary products were C_{2+} olefins and paraffins with an o/p ratio of 1.3 under typical FTS conditions ($H_2/CO = 2$, 300°C, 20 bar, 12,000 $ml \cdot g_{cat}^{-1} \cdot hour^{-1}$). The native iron carbide catalysts favored CO_2 formation and exhibited similar selectivity toward olefins and paraffins.

We then introduced 20 ppm of halomethanes (CH_3X, where X = F, Cl, Br, or I) into the syngas during FTS. CH_3F reduced CO_2 selectivity to 27.8%, accompanied by a slight increase in olefin content. A progressive trend was observed from CH_3F to CH_3Cl, CH_3Br, and CH_3I, with CO_2 selectivity continuously decreasing (Fig. 1A and table S1). Notably, cofeeding CH_3Br or CH_3I suppressed CO_2 formation, yielding near-zero CO_2 selectivity (~1%). Concurrently, olefin selectivity was substantially enhanced, with the o/p ratio rising to ~7 in the case of CH_3Cl. These results demonstrate that cofeeding trace amounts of halomethane effectively suppressed CO_2 production while promoting olefin selectivity in Fe-based FTS catalysts.

We next investigated the effect of CH_3Br concentration in the syngas on performance of FTS. Even a low cofeed concentration of 3 ppm CH_3Br markedly suppressed CO_2 formation, reducing its selectivity to 4.2%, decreasing CO conversion from 89.4% to 41.1%, and increasing the o/p ratio from 1.0 to 6.1 (Fig. 1B, figs. S1 and S2, and table S2). Chain propagation remained unaffected, and the olefin selectivity was much higher than that obtained with the unmodified χ-Fe_5C_2 catalyst (81% versus 23%; Fig. 1B, figs. S3 and S4, and table S2). The decrease in CO conversion mainly came from the inhibition of the CO_2 formation (fig. S2), which accounted for >70% of the reduction

[1]State Key Laboratory of Coal Conversion, Institute of Coal Chemistry, Chinese Academy of Sciences, Taiyuan, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]Beijing National Laboratory for Molecular Sciences, New Cornerstone Science Laboratory, College of Chemistry and Molecular Engineering, Peking University, Beijing, China. [4]College of Materials Science and Engineering, Beijing University of Technology, Beijing, China. [5]School of Chemistry and Chemical Engineering, Ningxia University, Yinchuan, China. [6]National Energy Center for Coal to Liquids, Huairou District, Beijing, China. [7]Max Planck Centre on the Fundamentals of Heterogeneous Catalysis (FUNCAT), Cardiff Catalysis Institute, School of Chemistry, Cardiff University, Cardiff, UK. ***Corresponding author. Email: liuxingwu@sxicc.ac.cn (X.-W.L.); wxd@sxicc.ac.cn (X.-D.W.); dma@pku.edu.cn (D.M.)** †These authors contributed equally to this work.

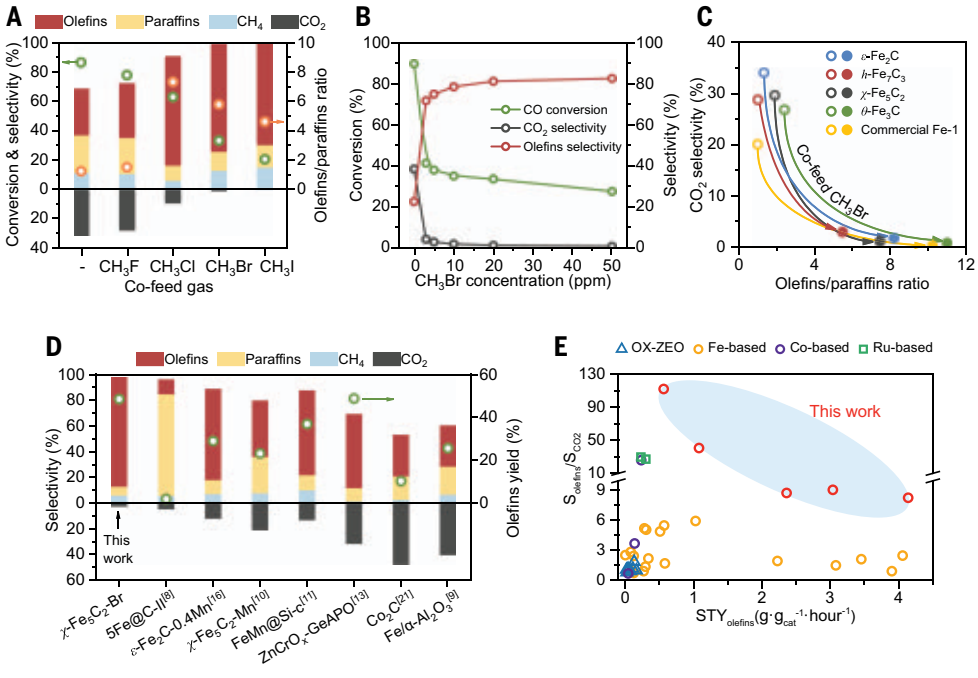


Fig. 1. Impact of halogen cofeeding on catalytic performance in FTS. (**A**) Effect of cofeeding halomethanes (CH_3F, CH_3Cl, CH_3Br, and CH_3I, 20 ppm) on CO conversion and product selectivity in FTS reaction over χ-Fe_5C_2 catalyst. Reaction conditions: 300°C, 20 bar, 12,000 ml·g_{cat}^{-1}·$hour^{-1}$, $H_2/CO = 2$. (**B**) Dependence of CO conversion and selectivity of CO_2 and olefin on the concentration of cofeeding CH_3Br (0 to 50 ppm). Reaction conditions: 300°C, 10 bar, 12,000 ml·g_{cat}^{-1}·$hour^{-1}$, $H_2/CO = 2$. (**C**) Influence of olefin/paraffin ratio and CO_2 selectivity for different iron carbide catalysts (ε-Fe_2C, *h*-Fe_7C_3, χ-Fe_5C_2, and θ-Fe_3C) and commercial Fe-1 under 20 ppm CH_3Br cofeeding. Reaction conditions: 300°C, 10 bar, $H_2/CO = 2$, CO conversion controlled at ~30% for all data points. (**D**) Comparison of catalytic selectivity and olefin yield of χ-Fe_5C_2 catalyst with CH_3Br cofeed against representative literature-reported catalysts. The χ-Fe_5C_2-Br catalyst demonstrates simultaneous ultralow CO_2 selectivity and higher olefin yield compared with state-of-the-art catalytic systems. (**E**) Plot of olefin-to-CO_2 selectivity ratio ($S_{olefins}/S_{CO2}$) against olefin space-time yield ($STY_{olefins}$) on χ-Fe_5C_2-Br (this work, red circles) and various catalytic systems from literature: OX-ZEO (*12*,*13*,*37*–*41*) (blue triangles), Fe-based (*8*–*11*,*16*,*20*,*27*,*32*,*42*–*45*) (yellow circles), Co-based (*21*,*46*,*47*) (purple circles), and Ru-based (*22*) (green squares). Note that the selectivity of olefins is based on all carbon-containing products. The shaded region highlights the exceptional catalytic performance of the χ-Fe_5C_2-Br catalyst.

in CO conversion, in the case of 3 ppm CH_3Br cofeeding (table S3). Although previous studies reported that cofeeding HCl or HBr can reduce CO_2 selectivity (*20*), they also observed a pronounced decline in FTS activity and limited CO_2 suppression (fig. S5). In contrast, CH_3Br cofeeding in this work enables near-zero CO_2 formation with minimal loss in FTS activity (fig. S5 and table S3). As the concentration of cofed CH_3Br increased, CO_2 selectivity progressively decreased—dropping to <1% at 50 ppm—whereas olefin selectivity showed a slight increase (Fig. 1B).

This halogen modulation strategy was further extended to other iron carbide phases, including ε-Fe_2C, θ-Fe_3C, and *h*-Fe_7C_3 (fig. S6) (*17*, *18*). In all cases, ppm concentrations of CH_3Br cofeeding suppressed CO_2 selectivity to <5% and enhanced o/p ratios by factors of three to four (Fig. 1C), underscoring the broad applicability and robustness of the approach. When we applied this strategy to a commercial iron catalyst ("Commercial Fe-1"), CO_2 selectivity could be suppressed to <1% (Fig. 1C).

To benchmark this halogen cofeeding method, we performed comparative analyses with state-of-the-art olefin synthesis catalysts, that is, Fe- and Co-based FTO catalysts and OX-ZEO systems (*8*–*11*, *13*, *16*, *21*) (Fig. 1D, figs. S7 to S11, and tables S4 to S8). Notably, the χ-Fe_5C_2 catalyst with Br cofeeding (denoted χ-Fe_5C_2-Br) exhibited superior olefin yield (~48% of total carbon-containing products) while maintaining ultralow CO_2 selectivity (<3%). In contrast, previously reported Fe- or Co-based and OX-ZEO catalysts typically exhibited CO_2 selectivity exceeding 10% or had a relatively low olefin yield (figs. S12 and S13).

Our strategy overcame the traditional trade-off between high olefin yield and low CO_2 selectivity, demonstrating unprecedented performance in the syngas-to-olefins reaction. A plot of the ratio of olefin-to-CO_2 selectivity ($S_{olefins}/S_{CO2}$) versus olefin space-time yield ($STY_{olefins}$) shows the overall catalytic performance (Fig. 1E). Although OX-ZEO catalysts offer good olefin selectivity among hydrocarbons, their low $STY_{olefins}$ and high CO_2 emissions constrain their practical utility (*13*). Similarly, Ru- and Co-based systems suffer from low $STY_{olefins}$ under conventional FTS conditions (*21*, *22*). In contrast, χ-Fe_5C_2-Br (Cl) catalysts delivered both high $STY_{olefins}$ and favorable olefin-to-CO_2 ratios. These results establish halogen cofeeding as an effective solution that enables high olefin space-time yield and high olefin yield with low CO_2 emissions in FTS. If C_{5+} olefins are used as liquid fuels, this strategy could also benefit conventional Fe-based FTS by blocking CO_2 emissions in standard FTS processes.

We also tracked the fate of CH_3Br in the product stream. Of the 20 ppm CH_3Br introduced, ~70% (14.2 ppm) was recovered as unreacted CH_3Br in the outlet gas, with <6 ppm detected in the aqueous phase. Only trace amounts of Br (<1 ppm) were found in the oil-phase products, indicating negligible incorporation into the hydrocarbon products (fig. S14 and tables S9 to S11). Notably, although only a small amount of Br remained on the surface of the iron catalyst (fig. S15 and table S12), it exerted a pronounced effect on the overall catalytic performance.

Structure of χ-Fe_5C_2 catalyst with CH_3Br cofeed

Cofeeding halomethanes induced a drastic shift in the reaction behavior of Fe-based FTS and largely suppressed CO_2 emissions (fig. S16). We hypothesized that halogen species play a critical role in modulating the structure of the working iron catalyst. To probe the structural consequences of halogen exposure, we characterized the χ-Fe_5C_2 catalyst after CH_3Br-cofeeding reaction (referred to as χ-Fe_5C_2-Br-spent). The catalysts with or without cofeeding halogen had similar morphologies and particle sizes (fig. S17). X-ray diffraction (XRD) analysis revealed that, in addition to the dominant χ-Fe_5C_2 phase, χ-Fe_5C_2-Br-spent exhibited new diffraction peaks associated with Fe_3O_4. In contrast, the control sample (χ-Fe_5C_2-spent, without CH_3Br) predominantly retained the carbide phase (Fig. 2A).

These results were corroborated by Mössbauer spectroscopy, electron microscopy, and extended x-ray absorption fine structure (EXAFS) fitting (figs. S18 to S24 and tables S13 to S16). X-ray photoelectron spectroscopy (XPS) further confirmed the presence of surface FeO_x, with detectable iron oxide species on both spent samples (Fig. 2B). The presence of FeO_x is commonly associated with the WGS reaction and therefore the CO_2 production in Fe-based FTS (*23*, *24*). Thus, the role of halogen was not to suppress the formation of iron oxides during the reaction.

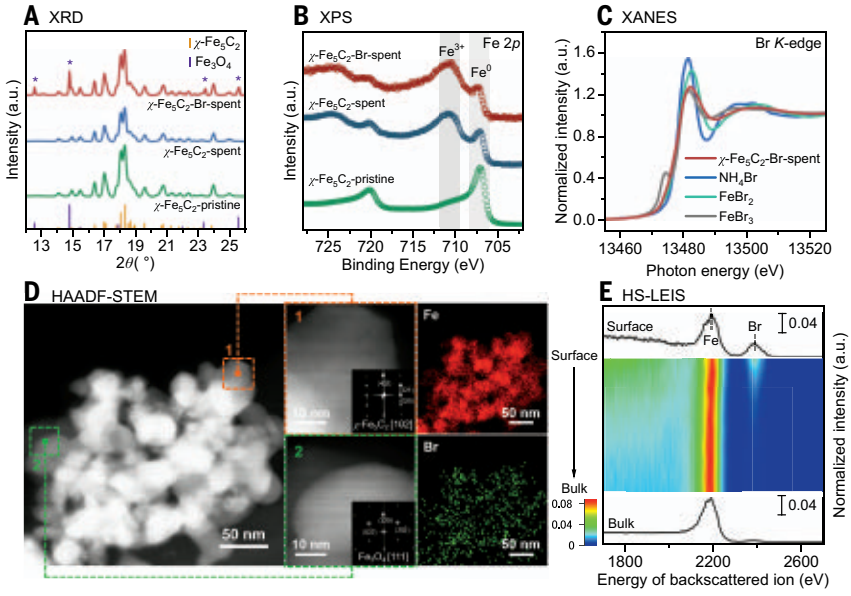


Fig. 2. Structure of χ-Fe_5C_2 and χ-Fe_5C_2-Br catalysts. Synchrotron XRD patterns (**A**) and Fe 2p XPS spectra (**B**) of pristine and spent χ-Fe_5C_2 and χ-Fe_5C_2-Br catalysts. a.u., arbitrary units. (**C**) Br K-edge XANES spectra of χ-Fe_5C_2-Br-spent catalyst compared with reference bromide compounds. (**D**) HAADF-STEM image and corresponding EDS mapping of Fe and Br distributions in χ-Fe_5C_2-Br-spent catalyst. (**E**) HS-LEIS depth profile of χ-Fe_5C_2-Br-spent catalyst. Reaction conditions: 100 mg catalyst, 300°C, 10 bar, 12,000 ml·g_{cat}^{-1}·$hour^{-1}$, $H_2/CO = 2$. χ-Fe_5C_2, syngas; χ-Fe_5C_2-Br, syngas with cofeeding 20 ppm CH_3Br. Catalysts were pretreated at 280°C and 1 bar for 6 hours, 12,000 ml·g_{cat}^{-1}·$hour^{-1}$, $H_2/CO = 2$ (χ-Fe_5C_2, syngas; χ-Fe_5C_2-Br, syngas with cofeeding 20 ppm CH_3Br).

Notably, elemental analysis, Br K-edge x-ray absorption near-edge structure (XANES), and Br 3d XPS revealed residual bromine species on χ-Fe_5C_2-Br-spent (Fig. 2C and fig. S25). The Br K-edge XANES spectra exhibited fingerprinting features similar to those of iron bromide, indicating strong Fe-Br interactions like that in iron bromide. High-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) imaging coupled with energy-dispersive x-ray spectroscopy (EDS) mapping showed homogeneous Br distribution across both Fe_3O_4 and χ-Fe_5C_2 domains (Fig. 2D and figs. S26 and S27), demonstrating homogeneous colocalization of Br and Fe species.

It remained uncertain whether Br was incorporated into the bulk lattice or was distributed on the surface. XRD showed no evidence of a new bromine-containing iron phase, suggesting that a surface-bound Br species was more likely (Fig. 2A and fig. S6). To confirm this hypothesis, we performed variable-energy XPS, which revealed substantial Br surface enrichment on the χ-Fe_5C_2-Br-spent catalyst (fig. S28). High-sensitivity low-energy ion scattering spectroscopy (HS-LEIS) is a surface-sensitive technique (*25*). From the HS-LEIS depth profile of the catalyst, we observed that Br was primarily enriched on the surface of the catalyst (Fig. 2E). Collectively, these findings strongly suggest that surface-bound Br modulates the chemical environment of active sites of Fe catalysts (both the iron carbide and oxide domains), suppressing CO_2 formation and enhancing olefin selectivity.

Inhibition mechanism of Br on CO_2 and paraffin formation

The reaction networks of the FTS process depicted in Fig. 3A were used to understand the mechanistic origin of this inhibition effect. Aside from the main reactions leading to the production of olefin and paraffin, there are two major side reactions that produce CO_2: WGS (highlighted in red) over iron oxide and the Boudouard reaction ($2CO \rightarrow C + CO_2$, highlighted in blue) over iron carbide (*14*, *26*). Although the primary products of Fe-based FTS were olefins, undesirable hydrogenation of olefins could lead to the production of less valuable paraffins ($C_nH_{2n} + H_2 \rightarrow C_nH_{2n+2}$, highlighted in purple) and a low o/p ratio (*27*).

On χ-Fe_5C_2 catalysts, both the Boudouard reaction and WGS routes contributed appreciably across typical reaction temperatures, with the WGS reaction becoming more pronounced at high conversions (fig. S29) (*15*, *24*, *26*, *28*). The drastic suppression of CO_2 formation observed in halogen-modified FTS suggested a fundamental alteration of reaction pathways by the Br on the surface of iron catalysts (fig. S30). To elucidate the underlying mechanism, we investigated catalytic reactivity by using transient kinetic analyses and surface-sensitive characterizations as well as computational modeling of reaction energetics, focusing on these two dominant CO_2-producing reactions in Fe-based FTS.

Both iron oxides (e.g., Fe_3O_4 and FeO_x clusters) and iron carbides may contribute to the observed WGS activity (*23*, *26*, *29*–*31*). In this study, Fe_3O_4 was selected as the model WGS catalyst under our conditions, while acknowledging that iron carbides can also play a role depending on temperature and surface state. We found that CH_3Br cofeeding could suppress WGS activity by more than two orders of magnitude (Fig. 3B and fig. S31). Given the established WGS mechanism wherein H_2O dissociates to generate surface hydroxyl (OH*), which then reacts with CO to form CO_2 (figs. S32 and S33), we first checked the adsorption and activation behavior of H_2O on the catalysts. Temperature-programmed desorption of water (H_2O-TPD) experiments revealed that both Fe_3O_4 and Br-pretreated Fe_3O_4 (referred to as Fe_3O_4-Br) showed similar H_2O adsorption abilities (desorbed at ~120°C). However, H_2 formation, an indicator for H_2O dissociation, was active on Fe_3O_4 but completely absent on Fe_3O_4-Br (Fig. 3C). In situ XPS further confirmed this suppression effect, showing minimal OH* production on Fe_3O_4-Br compared with Fe_3O_4 (Fig. 3D, figs. S34 and S35, and tables S17 and S18). Diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) and transient kinetic analysis (TKA) experiments corroborated this inhibition effect (figs. S33 and S36). These results established that Br disrupted the key step of H_2O dissociation in the WGS reaction, then effectively blocking CO_2 generation through this pathway.

These experimental findings were supported by DFT calculations on Fe_3O_4(111). In the absence of Br, H_2O can readily dissociate ($H_2O \rightarrow OH^* + H^*$) with the activation barrier (E_{a2}) of 0.73 eV. However, increasing Br coverage from $\theta_{Br} = 1/5$ to 1/3 raised this barrier to $E_{a2} >$ 1.2 eV. At a higher coverage of $\theta_{Br} = 2/5$, H_2O dissociation was completely inhibited (fig. S37). These findings aligned with TPD and XPS results, suggesting that surface-bound Br passivates reactive Fe sites and disrupts WGS reaction.

We then considered the Boudouard reaction. We performed CO pulse experiments on χ-Fe_5C_2 and χ-Fe_5C_2-Br catalysts. Both catalysts exhibited similar CO adsorption properties (figs. S38 and S39). However, CO_2 formation was observed on χ-Fe_5C_2, which was nearly completely suppressed on the χ-Fe_5C_2-Br in CO pulse, indicating that the Boudouard reaction was largely inhibited (Fig. 3E). Moreover, $Ar \rightarrow C^{18}O$ TKA experiments also showed no production of $C^{16}O^{18}O$ and $C^{18}O_2$ on χ-Fe_5C_2-Br but formation of $C^{16}O$ after introduction of $C^{18}O$, which may result from the recombination of dissociatively adsorbed CO (Fig. 3F). Notably, introduction of ^{13}CO to ^{18}O-precovered surface yielded $^{13}C^{18}O$ on χ-Fe_5C_2-Br (fig. S40), confirming that CO dissociation remained active. Temperature-programmed surface reaction (TPSR) experiments using a mixture of $C^{18}O$/He/Ar as feed further substantiated this selective blocking effect (fig. S41). These results suggest that

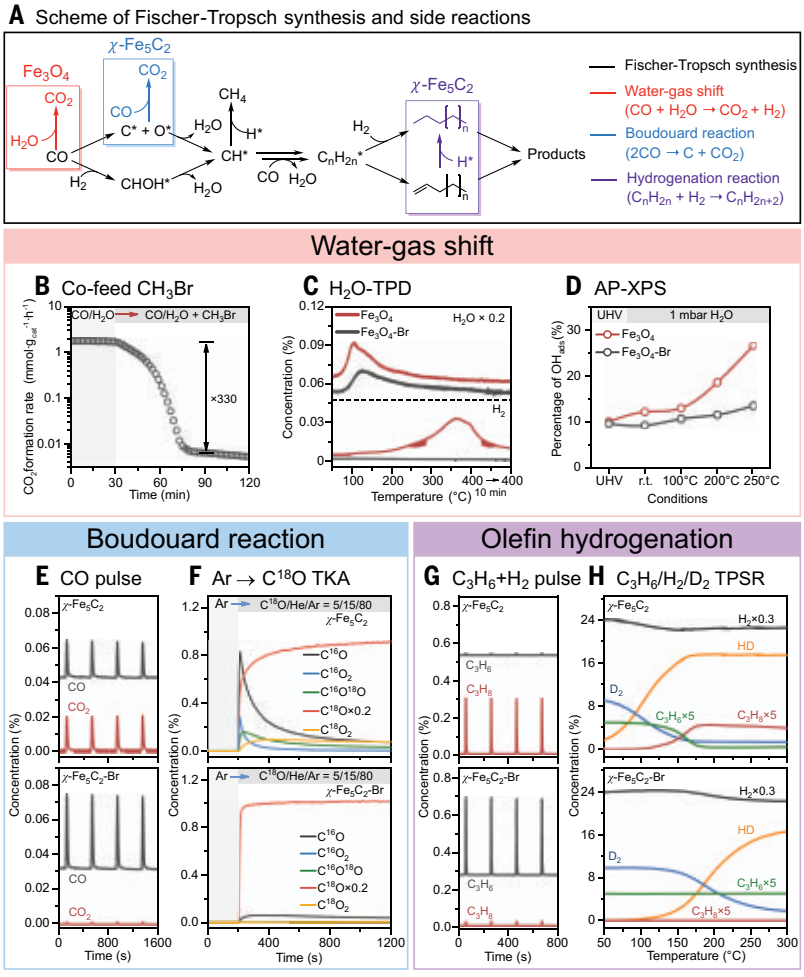


Fig. 3. Mechanism of Br in regulating catalytic behavior. (**A**) The reaction pathways in the FTS network. (**B**) Inhibition effect of CH_3Br cofeeding on CO_2 formation in water-gas shift reaction over Fe_3O_4 catalyst. Reaction conditions: 50 mg catalyst, 300°C, 1 bar, 10 ml/min $CO/H_2O/Ar = 2/2/96$ switched to $CO/H_2O/Ar = 2/2/96$ with cofeeding 300 ppm CH_3Br. (**C**) Temperature-programmed desorption of H_2O over Fe_3O_4 and Fe_3O_4-Br catalysts. Reaction conditions: 100 mg catalyst, 1 bar, pretreated in 20 ml/min $H_2O/Ar = 2.5/97.5$ at 50°C; then switched to 20 ml/min Ar, purging for 20 min, followed by temperature ramping from 50°C to 300°C at 10°C/min. (**D**) Near ambient pressure XPS results of Fe_3O_4 and Fe_3O_4-Br catalysts at different temperatures in 1 mbar H_2O. Conditions: Catalysts were pretreated at 400°C in Ar for 1 hour, cooled to room temperature, introduced to 1 mbar H_2O, and treated at different temperatures. UHV, ultrahigh vacuum. (**E**) CO pulse experiment over χ-Fe_5C_2 and χ-Fe_5C_2-Br catalysts. Reaction conditions: 100 mg catalyst, 300°C, 1 bar, 5 ml/min Ar, 250 μl sampling loop, pulse gas: CO/He = 0.5/99.5. (**F**) Transient kinetic analysis of $Ar \rightarrow C^{18}O/He/Ar = 5/15/80$ over χ-Fe_5C_2 and χ-Fe_5C_2-Br catalysts. Reaction conditions: 50 mg catalyst, 300°C, 1 bar, 10 ml/min Ar switched to $C^{18}O/He/Ar = 5/15/80$. (**G**) Pulse experiment of C_3H_6 in H_2 for χ-Fe_5C_2 and χ-Fe_5C_2-Br catalysts. Reaction conditions: 50 mg catalyst, 300°C, 1 bar, 20 ml/min H_2, 250 μl sampling loop, pulse gas: $C_3H_6/Ar/He = 3/3/94$. (**H**) Temperature-programmed experiment of C_3H_6 hydrogenation for χ-Fe_5C_2 and χ-Fe_5C_2-Br catalysts with cofeeding D_2. Reaction conditions: 50 mg catalyst, 1 bar, 10 ml/min $C_3H_6/H_2/D_2/Ar = 1/82/10/7$, temperature ramping from 50°C to 300°C at 10°C/min.

Br selectively inhibits the CO*-O* recombination step while preserving CO dissociation activity, which is in good agreement with DFT calculations (fig. S42).

From the results presented above, we conclude that surface Br species suppressed both WGS and Boudouard reactions. Another question is how Br modulated the o/p ratio in the reaction. Although promoters such as alkali metals or Mn for Fe-based catalysts are typically required to achieve a high olefin selectivity (*16*, *27*, *32*, *33*), we added no alkali metal or Mn in the current high-olefin-selectivity χ-Fe_5C_2-Br catalysts. To understand the origin of the ability of χ-Fe_5C_2-Br to achieve a high o/p ratio, we used propene hydrogenation as a probe to study the olefin hydrogenation process. Propene pulse experiments in H_2 showed a drastic reduction in propene hydrogenation activity on χ-Fe_5C_2-Br compared with χ-Fe_5C_2 (Fig. 3G), which indicated that surface Br species suppressed the olefin hydrogenation reaction.

Weakened H_2 activation ability on the Br-modified surface could be one reason for high olefin selectivity, but this explanation appeared unlikely, because hydrocarbon formation was still observed. Another possibility is that the olefin hydrogenation step is suppressed. To further substantiate this point, we conducted TPSR experiments using a $C_3H_6/H_2/D_2/Ar = 1/82/10/7$ mixture feed. Hydrogen deuteride (HD) formation through H-D exchange was used as a probe to monitor H_2 activation. On χ-Fe_5C_2, both H-D exchange and C_3H_6 hydrogenation increased with temperature, reaching near-complete C_3H_6 conversion at around 180°C. In contrast, on χ-Fe_5C_2-Br, although activation of H_2 was slightly suppressed at low temperatures and became very active above 200°C, no detectable hydrogenation of C_3H_6 was observed even at temperatures up to 300°C (Fig. 3H). These results demonstrated that the high olefin selectivity came from inhibition of olefin hydrogenation ability by Br.

DFT simulations confirmed that Br altered hydrogenation energetics. On clean χ-Fe_5C_2(100), the two-step hydrogenation of C_3H_6 proceeds with low barriers (0.22 and 0.26 eV), facilitating paraffin formation (Fig. 4A). With $\theta_{Br} = 1/3$, the barrier of the first step ($C_3H_6^* + H^* \rightarrow C_3H_7^*$) increased to 0.66 eV, surpassing C_3H_6 desorption energy, making desorption more favorable than hydrogenation. Detailed electronic analyses (figs. S43 to S45) and optimized transition state geometries (figs. S46 and S47) revealed that surface-bound Br species not only electronically modified Fe active sites but also introduced substantial steric hindrance. This steric effect obstructed the H* approach and altered the Fe-C_3H_6 binding geometry. As a result, secondary hydrogenation was selectively suppressed under Br-modified conditions.

Finally, we examined the catalytic stability of the χ-Fe_5C_2 catalyst with 20 ppm CH_3Br cofeeding, a critical measurement for practical applications. As shown in Fig. 4B, the catalyst exhibited excellent stability for >450 hours. At a steady CO conversion of ~35%, the system maintained an exceptionally high C_{2+} olefin selectivity of >80%. Notably, the formation of undesired by-products such as methane and CO_2 remained minimal throughout the test, with CO_2 selectivity consistently <1.5%. Such sustained performance—combining high olefin selectivity and near-zero CO_2 emissions—has not been previously reported in the century-long development of Fe-based FTS. These findings underscore the transformative potential of halogen cofeeding for enabling low-emission, high-efficiency olefin production at scale.

Conclusions

We demonstrated a simple yet powerful strategy to enhance FTS-to-olefins performance by cofeeding ppm levels of CH_3Br into Fe-based catalytic systems. Surface-bound Br species electronically interacted with Fe active sites, selectively inhibiting H_2O dissociation, CO*-O* recombination, and olefin hydrogenation, while largely preserving CO and H_2 adsorption and activation. The strategy reported in this work enables a near-zero-CO_2 pathway for highly selective olefin (or liquid fuels) production in Fe-based FTS processes that achieve high

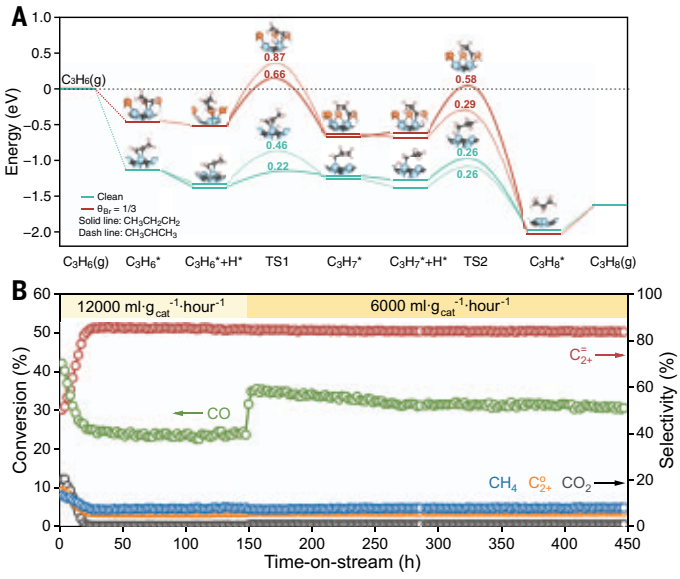


Fig. 4. Theoretical investigation of olefin hydrogenation inhibition mechanism and catalytic stability of χ-Fe_5C_2 catalyst with CH_3Br cofeed in the FTS process. (**A**) DFT-calculated reaction energy profiles for propene hydrogenation on clean and Br-covered ($\theta_{Br} = 1/3$) χ-Fe_5C_2(100) surfaces. Reaction intermediates and transition states illustrate two hydrogenation pathways, through $CH_3CH_2CH_2$ (solid line) and CH_3CHCH_3 (dotted line), highlighting increased hydrogenation barriers upon Br modification. (**B**) Catalytic stability and product selectivity of the χ-Fe_5C_2-Br catalyst. Reaction conditions: 100 mg catalyst, 300°C, 5 bar, $H_2/CO = 2$, cofeeding 20 ppm CH_3Br. The catalyst exhibited stable CO conversion and sustained high selectivity toward C_{2+} olefins with minimal CH_4 and CO_2 by-product formation over 450 hours at two different space velocities (12,000 and 6000 ml·g_{cat}^{-1}·$hour^{-1}$).

α-olefin selectivity and maintain industrially competitive activity, surpassing other reported low-CO_2 systems (table S8).

Although halogen-containing cofeed could pose challenges in scaled-up industrial reactors, similar strategies have been adopted in commercialized processes, such as chlorine cofeed silver-catalyzed ethylene epoxidation process (*34*, *35*). Coupled with CO_2-free gasification reaction (e.g., $CH_4 + 3CO_2 = 4CO + 2H_2O$) and green hydrogen from water electrolysis (*36*), the process developed in this work could offer a viable pathway toward carbon-neutral coal-to-liquid/olefins or gas-to-liquid/olefins processes.

REFERENCES AND NOTES

1. K. T. Rommens, M. Saeys, *Chem. Rev.* **123**, 5798–5858 (2023).
2. X. Pan, F. Jiao, D. Miao, X. Bao, *Chem. Rev.* **121**, 6588–6609 (2021).
3. J. Li *et al.*, *Nat. Catal.* **1**, 787–793 (2018).
4. H. Zhao *et al.*, *CCS Chem.* **3**, 2712–2724 (2021).
5. P. Wang *et al.*, *Sci. Adv.* **4**, eaau2947 (2018).
6. S. Lyu *et al.*, *Nat. Commun.* **11**, 6219 (2020).
7. D. Wang *et al.*, *ACS Catal.* **15**, 7028–7039 (2025).
8. X. Zhang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2407624121 (2024).
9. H. M. Torres Galvis *et al.*, *Science* **335**, 835–838 (2012).
10. P. Wang *et al.*, *Nature* **635**, 102–107 (2024).
11. Y. Xu *et al.*, *Science* **371**, 610–613 (2021).
12. F. Jiao *et al.*, *Science* **351**, 1065–1068 (2016).
13. F. Jiao *et al.*, *Science* **380**, 727–730 (2023).
14. E. de Smit *et al.*, *Angew. Chem. Int. Ed.* **50**, 1584–1588 (2011).
15. M. Ojeda *et al.*, *J. Catal.* **272**, 287–297 (2010).
16. F. Qian *et al.*, *Nat. Commun.* **15**, 5128 (2024).
17. X.-W. Liu *et al.*, *J. Phys. Chem. C* **121**, 21390–21396 (2017).
18. X. W. Liu *et al.*, *Sci. Rep.* **6**, 26184 (2016).
19. C. Yang, H. Zhao, Y. Hou, D. Ma, *J. Am. Chem. Soc.* **134**, 15814–15821 (2012).
20. W. Ma *et al.*, *ACS Catal.* **5**, 3124–3136 (2015).
21. L. Zhong *et al.*, *Nature* **538**, 84–87 (2016).
22. H. Yu *et al.*, *Nat. Commun.* **13**, 5987 (2022).
23. M. Claeys *et al.*, *ACS Catal.* **11**, 13866–13879 (2021).
24. E. de Smit *et al.*, *J. Am. Chem. Soc.* **132**, 14928–14941 (2010).
25. C. Zhan *et al.*, *J. Am. Chem. Soc.* **142**, 14134–14141 (2020).
26. C. J. Keturakis *et al.*, *ACS Catal.* **6**, 4786–4798 (2016).
27. P. Zhai *et al.*, *Angew. Chem. Int. Ed.* **55**, 9902–9907 (2016).
28. R. Gubo *et al.*, *ACS Catal.* **14**, 14721–14732 (2024).
29. M. Zhu, I. E. Wachs, *ACS Catal.* **6**, 722–732 (2015).
30. F. Polo-Garzon *et al.*, *J. Am. Chem. Soc.* **141**, 7990–7999 (2019).
31. H. Yan *et al.*, *Appl. Catal. B* **226**, 182–193 (2018).
32. Y. Liu, J.-F. Chen, J. Bao, Y. Zhang, *ACS Catal.* **5**, 3905–3909 (2015).
33. R. Wang *et al.*, *ACS Catal.* **14**, 11121–11130 (2024).
34. V. P. Santos *et al.*, *ACS Catal.* **14**, 10839–10852 (2024).
35. R. M. Lambert, R. L. Cropley, A. Husain, M. S. Tikhov, *Chem. Commun.* **2003**, 1184–1185 (2003).
36. L. C. Buelens, V. V. Galvita, H. Poelman, C. Detavernier, G. B. Marin, *Science* **354**, 449–452 (2016).
37. K. Cheng *et al.*, *Angew. Chem. Int. Ed.* **55**, 4725–4728 (2016).
38. F. Jiao *et al.*, *Angew. Chem. Int. Ed.* **57**, 4692–4696 (2018).
39. Y. Ni *et al.*, *ACS Catal.* **9**, 1026–1032 (2019).
40. Y. Zhu *et al.*, *ACS Catal.* **7**, 2800–2804 (2017).
41. J. Su *et al.*, *Nat. Commun.* **10**, 1297 (2019).
42. J. Lu *et al.*, *ACS Catal.* **4**, 613–621 (2014).
43. O. Zhuo *et al.*, *Chem. Sci.* **10**, 6083–6090 (2019).
44. Y. He, H. Shi, O. Johnson, B. Joseph, J. N. Kuhn, *ACS Catal.* **11**, 15177–15186 (2021).
45. J. Wang *et al.*, *ACS Appl. Mater. Interfaces* **10**, 43578–43587 (2018).
46. J. Xie *et al.*, *Nat. Commun.* **10**, 167 (2019).
47. T. Lin *et al.*, *Appl. Catal. B* **299**, 120683 (2021).

ACKNOWLEDGMENTS

This work is dedicated to the 100th anniversary of the FTS process. The authors acknowledge the staff of the BL02U2 beamlines at the Shanghai Synchrotron Radiation Facility (https://cstr.cn/31124.02.SSRF.BL02U2) for their technical support and assistance with XRD measurements. The authors acknowledge the BL02B01 and BL20U1 beamlines at the Shanghai Synchrotron Radiation Facility (https://cstr.cn/31124.02.SSRF.BL02B01, https://cstr.cn/31124.02.SSRF.BL20U1) for technical support and assistance with XPS and XAS measurements, respectively. The authors thank the BL01B1 beamline at SPring-8, Japan, for their valuable technical support and assistance with XAS data collection. The artificial intelligence (AI)–driven experiments, simulations, and model training were performed on the robotic AI-Scientist platform of the Chinese Academy of Sciences. We thank W. Zhou and J. Gao for helpful discussions. The authors gratefully acknowledge the resources and financial support provided by the Research and Design Center of Synfuels China Co., Ltd. D.M. acknowledges support from the New Cornerstone Science Foundation through the XPLORER PRIZE and New Cornerstone Investigator Program. **Funding:** National Science Fund for Distinguished Young Scholars of China 22225206 (X.-D.W.) and 22025804 (Y.Y.); National Key R&D Program of China 2022YFB4101200 (Y.Y.), 2023YFB4103100 (Y.X.), 2021YFA1501102 (D.M.), and 2021YFA1500300 (X.L.); National Natural Science Foundation of China 22172183 (X.-W.L.), 22572210 (X.-W.L.), 22232001 (D.M.), and 22272106 (X.L.); Beijing Natural Science Foundation L255001 (X.-W.L.); Ordos Key R&D Program YF20232316 (X.-W.L.) and YF20232317 (X.-D.W.). **Author contributions:** D.M., X.-D.W., and X.-W.L. conceived of and supervised the project. Y.C. and X.-W.L. conducted most of the catalytic reactions. M.W. performed the transient kinetic experiments. F.Q. performed Mössbauer spectroscopy. M.W., L.Z., Z.C., and Y.X. carried out and analyzed the x-ray–based characterizations (XRD, XAS, and XPS). Y.C., X.-W.L., and X.L. conducted the electron microscopy studies. S.Z., X.-W.L., and X.-D.W. performed the DFT calculations. Y.C., M.W., X.-W.L., S.Z., and D.M. wrote the original draft. B.X., G.J.H., H.W., J.X., J.Z., Y.-W.L., Y.Y., X.L., X.C.L., and Z.R.G. provided scientific guidance and critical discussion. All authors contributed to discussions and revisions of the manuscript. **Competing interests:** A Chinese patent covering the strategy of cofeeding halogen-containing gases, as reported in this paper, has been filed. Other than this, the authors declare that they have no competing interests. **Data and materials availability:** All data supporting the findings of this study are available within the paper and its supplementary materials. **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.aea1655
Materials and Methods; Figs. S1 to S47; Tables S1 to S18; References (*48*–*57*)

Submitted 28 June 2025; accepted 8 September 2025

10.1126/science.aea1655

COMPARATIVE COGNITION

Chimpanzees rationally revise their beliefs

Hanna Schleihauf[1,2]†, Emily M. Sanford[3]†, Bill D. Thompson[3], Snow Zhang[4], Joshua Rukundo[5], Josep Call[6,7], Esther Herrmann[8]‡, Jan M. Engelmann[3]*‡

The selective revision of beliefs in light of new evidence has been considered one of the hallmarks of human-level rationality. However, tests of this ability in other species are lacking. We examined whether and how chimpanzees (*Pan troglodytes*) update their initial belief about the location of a reward in response to conflicting evidence. Chimpanzees responded to counterevidence in ways predicted by a formal model of rational belief revision: They remained committed to their initial belief when the evidence supporting the alternative belief was weaker, but they revised their initial belief when the supporting evidence was stronger. Results suggest that this pattern of belief revision was guided by the explicit representation and weighing of evidence. Taken together, these findings indicate that chimpanzees metacognitively evaluate conflicting pieces of evidence within a reflective process.

Rationality is a hallmark of human thinking. Rational agents form beliefs about the world on the basis of evidence (*1*), but as new evidence emerges, they revise their beliefs by weighing the strength of their prior beliefs in relation to the strength of the new evidence (*2*). Selective belief revision is a particularly strong indicator of rationality (*3*).

From a Bayesian perspective, the rational response to counter-evidence is to revise a belief in hypothesis h_1 to a belief that supports an alternative hypothesis h_2 when the counterevidence causes the posterior probability of h_2 to exceed the posterior probability of h_1:

$$\frac{p(h_2 \mid e_{Old}, e_{New})}{p(h_1 \mid e_{Old}, e_{New})} = \frac{p(e_{Old}, e_{New} \mid h_2)\, p(h_2)}{p(e_{Old}, e_{New} \mid h_1)\, p(h_1)} > 1$$

When both hypotheses have equal probabilities a priori, this condition is met when the strength of the evidence (e_{Old}) supporting the initial hypothesis is exceeded by the strength of new evidence (e_{New}) supporting the alternative hypothesis [see supplementary materials (SM)]. A reasoner should maintain their initial belief when the evidence in support of their initial belief is stronger and revise their belief if the evidence in support of the alternative belief is stronger.

Humans' closest primate relatives adjust their beliefs in response to evidence. For example, great apes draw inferences based on auditory, visual, and tactile cues during problem solving (*4*–*12*). Chimpanzees infer food locations from left-behind traces (*13*, *14*), weight effects (*15*), or a food's absence from an alternative location (*16*, *17*). They flexibly distinguish between appearance and reality, discounting misleading cues such as apparent food size (*18*), and they actively seek additional information when faced with contradictory evidence (*19*). It is not known, however, whether chimpanzees demonstrate the metacognitive ability to rationally evaluate the strength of new evidence and compare it with the strength of their existing beliefs—and so revise their beliefs only when appropriate, as humans can.

We tested whether chimpanzees' patterns of belief change conform to the predictions of a Bayesian formulation of rational belief revision (Experiments 1 and 2). We then probed the cognitive mechanisms underlying chimpanzees' belief revision. We asked whether chimpanzees represent both possible hypotheses before choosing between them (Experiment 3), individuate evidence (differentiate new from redundant evidence; Experiment 4), and respond to second-order evidence (evidence about the validity of the first-order evidence; Experiment 5).

Chimpanzees' response to counterevidence

In Experiments 1 and 2 (preregistered), chimpanzees ($N = 15$) were presented with a piece of food, which was then hidden in one of two possible locations. Next, they were given evidence that the food was in one of the locations and made their first choice. Subjects were then given a different form of evidence that favored the other location and made a second choice (which confirmed or changed their first choice). In the "strong evidence first" condition, chimpanzees were provided with strong evidence for one of the locations before their first choice and then weak evidence for the alternative location before their second choice. In the "weak evidence first" condition, the order was reversed (Fig. 1). Our classifications for strong versus weak evidence were based on previous research (*13*, *16*). In Experiment 1, we classified visual evidence (seeing food) as strong evidence and auditory evidence (hearing food) as weak evidence (Fig. 1A). In Experiment 2, we classified auditory evidence as strong evidence and indirect visual evidence (seeing food traces) as weak evidence (Fig. 1B). Chimpanzees' first choices validated these classifications (Experiment 1: $\chi^2 = 8.37$, df = 1, $p = 0.004$; Experiment 2: $\chi^2 = 11.99$, df = 1, $p = 0.001$; Fig. 2A and SM).

We assessed whether chimpanzees' choices aligned with the predictions of rational belief revision. In Experiment 1, chimpanzees were significantly more likely to revise their beliefs in the "weak evidence first" condition than in the "strong evidence first" condition ($\chi^2 = 14.03$, df = 1, $p < 0.001$; Fig. 2B). We found no evidence that chimpanzees simply learned the appropriate behavior over time: Neither trial number ($\chi^2 = 0.85$, df = 1, $p = 0.398$) nor the interaction between trial number and condition ($\chi^2 = 3.56$, df = 1, $p = 0.151$) had a significant effect on performance. In Experiment 2, chimpanzees also revised their belief significantly more often in the "weak evidence first" condition than in the "strong evidence first" condition ($\chi^2 = 20.52$, df = 1, $p < 0.001$; Fig. 2B). Again, we found no evidence that chimpanzees simply learned the behavior across trials: Neither trial number ($\chi^2 = 0.93$, df = 1, $p = 0.336$) nor the interaction between trial number and condition ($\chi^2 = 1.60$, df = 1, $p = 0.206$) had a significant effect on the likelihood of belief revision.

The results also confirm that chimpanzees' choices were in line with the principle of commutativity, which holds that one's belief should depend on the sum total of all previously perceived evidence, independent of the order in which that evidence was acquired (*20*). After receiving both strong and weak evidence, chimpanzees chose the location supported by strong evidence at above chance levels in both conditions in Experiment 1 [proportion choosing strong evidence side: weak evidence first condition = 0.93, 95% confidence interval (CI) (0.88, 1.00); strong evidence first condition = 0.92, 95% CI (0.85, 1.00)] and Experiment 2 [weak evidence first condition = 0.96, 95% CI (0.93, 1.00); strong evidence first condition = 0.86, 95% CI (0.77, 1.00)]. Moreover, the likelihood of endorsing the strong evidence at the second choice did not differ depending on whether it was presented first (strong evidence first condition) or second (weak evidence first condition) in Experiment 1 ($\chi^2 = 1.64$, df = 1, $p = 0.286$) or Experiment 2 ($\chi^2 = 1.40$, df = 1, $p = 0.330$).

[1]Department for Developmental Psychology, Utrecht University, Utrecht, Netherlands. [2]German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany. [3]Department of Psychology, University of California, Berkeley, Berkeley, CA, USA. [4]Department of Philosophy, University of California, Berkeley, Berkeley, CA, USA. [5]Ngamba Island Chimpanzee Sanctuary/Chimpanzee Trust, Entebbe, Uganda. [6]School of Psychology and Neuroscience, University of St Andrews, St. Andrews, UK. [7]Global Research Centre for Diverse Intelligences, University of St Andrews, St. Andrews, UK. [8]School of Psychology, Sport and Health Sciences, University of Portsmouth, Portsmouth, UK. ***Corresponding author. Email: jan_engelmann@berkeley.edu** †These authors contributed equally to this work. ‡These authors contributed equally to this work.

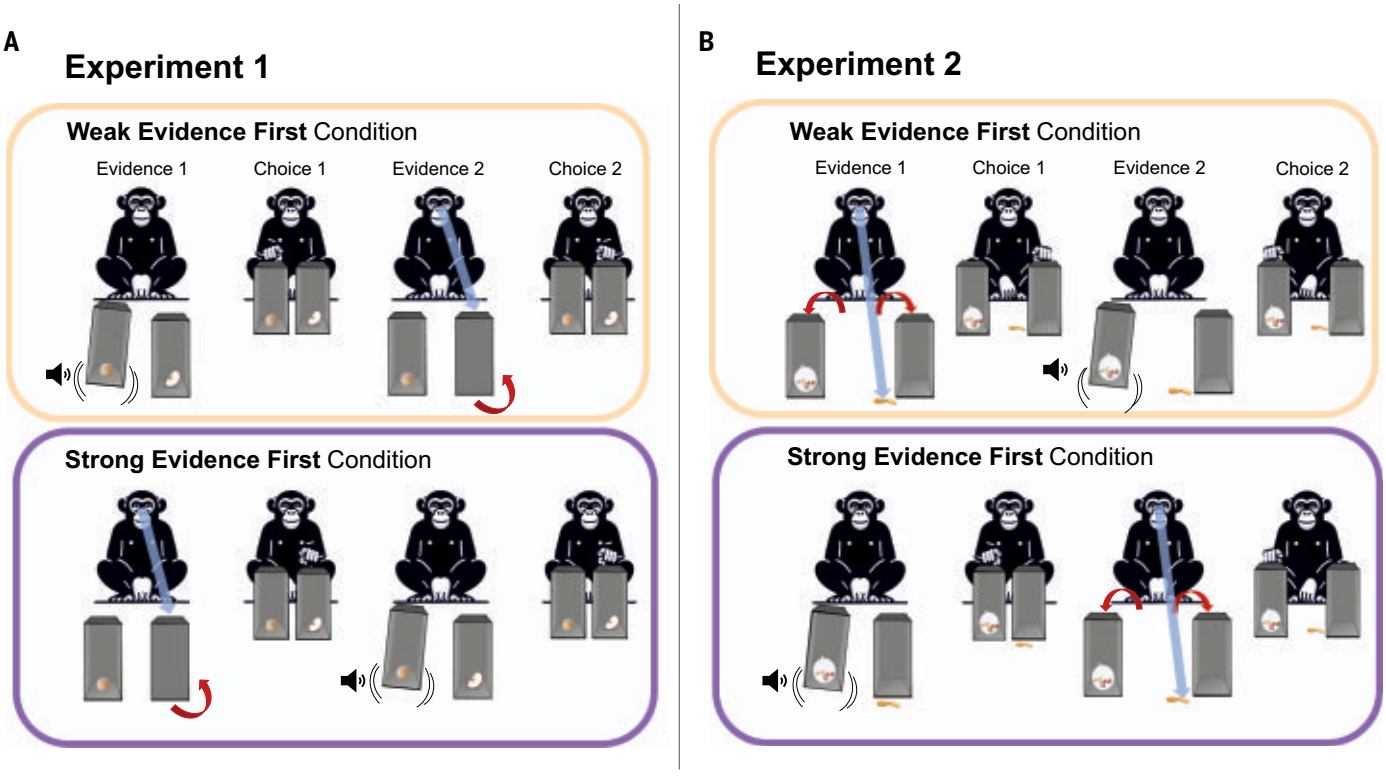


Fig. 1. Procedure of Experiments 1 and 2. (**A**) Experiment 1, Weak evidence first condition: Chimpanzees first received weak evidence (auditory evidence: shaking the box with a piece of wood in it) for one of the boxes and made a first choice. Then they received strong evidence (direct visual evidence: seeing into the box through a glass pane) for the other box and made a second choice. Strong evidence first condition: The same types of evidence were presented but in reverse order. (**B**) Experiment 2, Weak evidence first condition: Chimpanzees first received weak evidence (indirect visual evidence: seeing traces of the food behind the box) for one of the boxes and made a first choice. Then they received strong evidence (auditory evidence: shaking the box with a plastic capsule containing peanuts in it) for the other box and made a second choice. Strong evidence first condition: The same types of evidence were presented but in opposite order. Note: The red arrows indicate that the box was rotated (Experiment 1) or moved to the side (Experiment 2) to reveal visual evidence. In Experiments 1 and 2, in addition to the weak evidence first condition and the strong evidence first condition, chimpanzees also participated in filler trials, during which they received one piece of evidence, made a single choice, and received the contents of the box that they picked on their first choice (there was no second piece of evidence and no second choice). Filler trials were included such that subjects could not, on any given trial, predict whether their first or second choice would count (for details, see SM).

We also implemented a Bayesian modeling analysis in Stan (*21*) to assess whether chimpanzees' choice behavior is consistent with a formal model of belief revision (rational choice model; fig. S3 and SM). We found that the core conditions for rationality were met: Inferred evidence strength for strong evidence was both larger than—and in the opposite direction of—weak evidence (Fig. 2C and SM). Furthermore, this rationality did not merely emerge at the group level: Across both experiments, this pattern held for nearly every individual subject (Fig. 2, C and D, and SM).

To assess whether the rational choice model provided the best explanation for chimpanzees' choices, we compared it against four alternatives (fig. S3 and SM): belief perseveration (*22*–*24*), recency bias (*25*, *26*), cue saliency (*27*, *28*), and a null model. For both experiments, the rational choice model outperformed the recency bias, belief perseveration, and null models, while the cue saliency model performed equally well compared with the rational choice model (see SM).

The cognitive underpinnings of chimpanzees' belief revision

The findings of Experiments 1 and 2 suggest that chimpanzees compare evidence in favor of two alternative options and endorse the option supported by stronger evidence. It is possible, however, that receiving strong evidence—such as seeing the food directly in Experiment 1—simply replaced representations based on weak evidence, without any comparison across represented options (cue saliency model). In Experiment 3, we tested whether the saliency of the strongest cue overrides other representations. Subjects ($N = 23$) were presented with three hiding locations and received strong evidence (seeing food) for one of the locations, weak evidence (hearing food) for a second location, and no evidence for the third location (Fig. 3). Then, the box supported by strong evidence was removed. Lastly, chimpanzees made a choice. We tested whether chimpanzees would select the option supported by weak evidence more often than the option supported by no evidence, which would suggest that weaker representations are maintained even when the evidence supporting them is not the most salient. Alternatively, they could choose randomly between the two remaining options, suggesting that only the representation supported by the most salient evidence is maintained. Chimpanzees chose the weak evidence option at a rate significantly above chance ($\beta_0 = 1.41$, standard error $= 0.19$, $z = 7.61$, $p < 0.001$), and the Bayesian analysis confirmed that their weight on weak evidence was significantly greater than 0 (see SM). This suggests that chimpanzees indeed represented both options during Experiments 1 and 2 [see fig. S4 and SM for an additional experiment replicating this finding; see also (*29*–*32*)].

If chimpanzees in Experiments 1 and 2 explicitly represented the evidence in favor of both options and compared their relative strength, they should demonstrate two further capacities: differentiating between new and redundant evidence (Experiment 4) and sensitivity to second-order evidence—evidence about the validity of the previous evidence (Experiment 5).

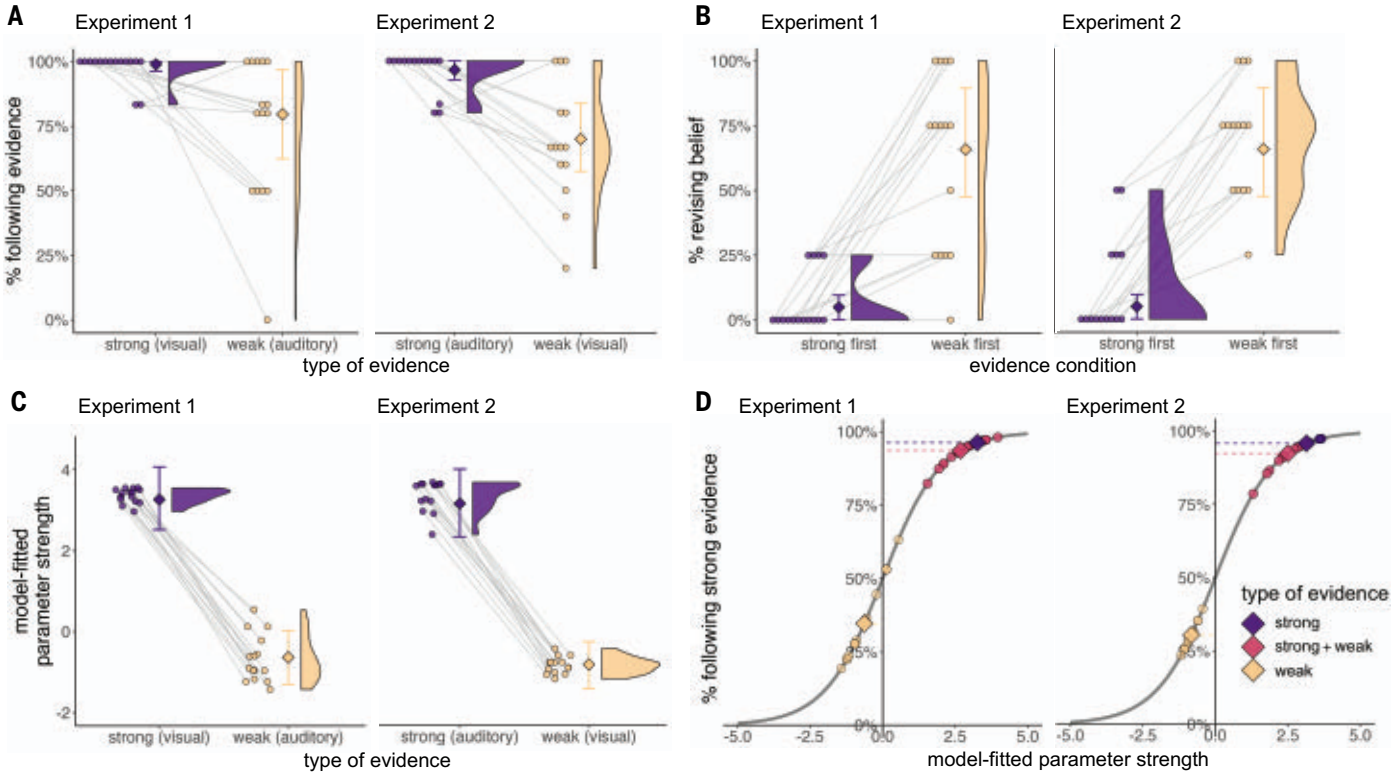


Fig. 2. Results of Experiments 1 and 2. (**A**) First choices by evidence type. Each point represents one chimpanzee's average response for that trial type. Error bars represent the bootstrapped 95% CIs of the generalized linear mixed model (GLMM) including all main effects. Diamonds represent the model predictions. Gray lines connect the behavior of one chimpanzee across both conditions within each experiment. (**B**) Belief revision by condition. Error bars represent the bootstrapped 95% CIs of the GLMM including all main effects. Diamonds represent the model predictions. Gray lines connect the behavior of one chimpanzee across both conditions within each experiment. (**C**) Rational choice model parameter estimates for individual subjects' e_{Strong} and e_{Weak} in Experiments 1 and 2. Error bars represent the model's 95% CIs on the group mean estimates (diamonds). (**D**) Predicted choice rates based on model-fitted evidence strengths across different evidence types (strong, weak, and strong + weak). Dots reflect individual subject parameters, and diamonds reflect group means. The dashed lines reflect the model's assumed relationship between evidence strength and choice rate.

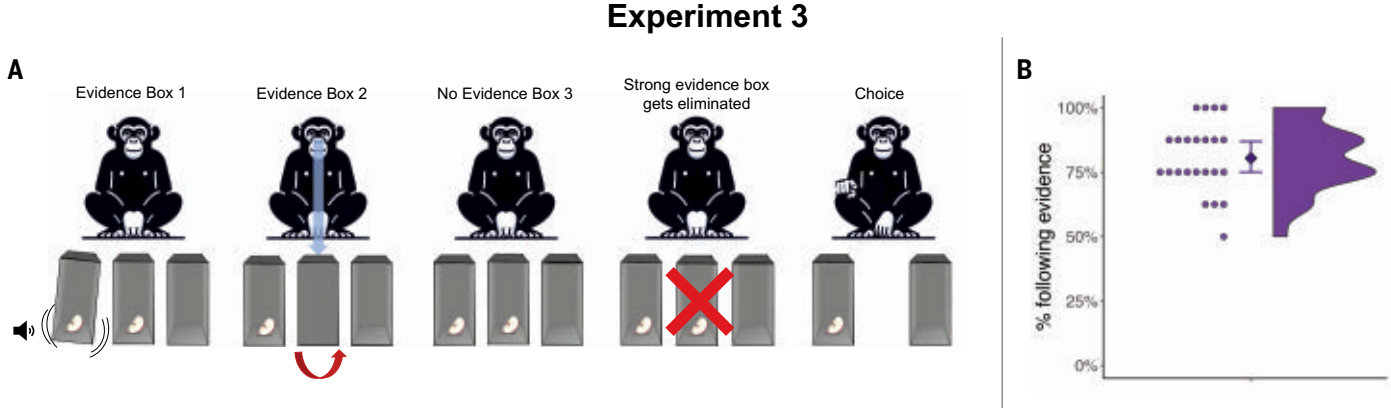


Fig. 3. Procedure and results of Experiment 3. (**A**) Chimpanzees received, in counterbalanced order, weak evidence for one box (auditory evidence: shaking the box with a piece of wood in it), strong evidence for a second box (direct visual evidence: seeing into the box through a glass pane), and no evidence for the third box (the box was touched for the same amount of time that each of the other two boxes was manipulated, but no evidence was provided). Then the experimenter removed the box supported by strong evidence before the chimpanzee was allowed to make a choice. (**B**) Chimpanzees were significantly more likely to select the box supported by weak evidence (rather than the box supported by no evidence) than would be expected by chance.

In Experiment 4, we asked whether chimpanzees ($N = 23$) would be swayed more by new evidence (evidence that comes from an independent, distinct source that provides additional information) compared with redundant evidence (evidence that comes from the same source as previous evidence and therefore provides no new information). In the "redundant evidence" condition, the weak evidence was auditory shaking evidence (which, when repeated, provides no new information), whereas in the "new evidence" condition, the weak evidence was auditory dropping evidence (which, when repeated, indicates that a second piece of food has been dropped into the box; Fig. 4 and SM).

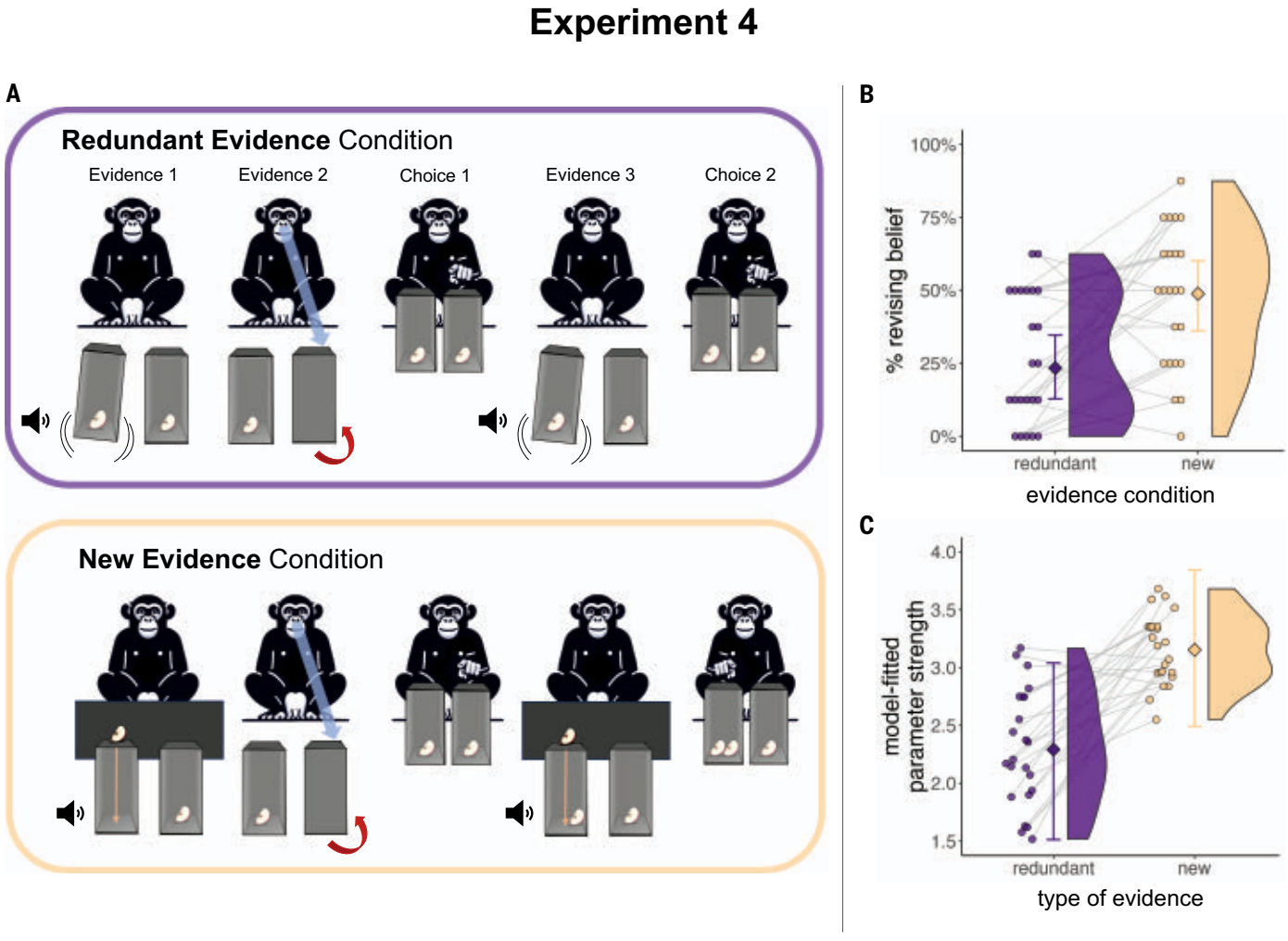


Fig. 4. Procedure and results of Experiment 4. Chimpanzees received weak evidence for one of the boxes, followed by strong evidence for the other box. They then made a first choice, after which they were presented with the same weak evidence again. The nature of the weak evidence varied according to condition. (**A**) In the redundant evidence condition, the weak evidence did not accumulate: The box was shaken a second time, providing no new information about its contents. In the new evidence condition, the weak evidence accumulated: The chimpanzees could infer on the basis of the second dropping noise that a second piece of apple had been dropped into the box. Notably, before presenting the second piece of evidence, the experimenter showed the chimpanzees a second piece of apple in both conditions. In the redundant evidence condition, the experimenter presented the apple and then placed it in their pocket. In the new evidence condition, the experimenter concealed the apple in their hands before dropping it into one of the boxes. (**B**) Chimpanzees were more likely to revise their beliefs when they heard a second piece of food drop into the box (new evidence condition) than when they heard the box shake for a second time (redundant evidence condition). (**C**) The rational choice model confirmed that for nearly every subject, the second piece of new evidence had a greater impact on their belief than the second piece of redundant evidence.

Across both conditions, chimpanzees were first presented with weak and strong (visual) evidence. Then, after they chose the option supported by the strong evidence, they were provided with a second instance of the weak evidence. Chimpanzees demonstrated a significantly higher tendency to revise their initial choice in the "new evidence" condition compared with the "redundant evidence" condition ($\chi^2 = 9.25$, df = 1, $p = 0.002$; Fig. 4). Their tendency to revise beliefs decreased over trials ($\chi^2 = 5.06$, df = 1, $p = 0.024$), with a steeper decline in the "redundant evidence" condition ($\chi^2 = 5.59$, df = 1, $p = 0.018$; fig. S5 and SM). When presented on its own, the auditory dropping evidence presented in the "new evidence" condition was weaker than the auditory shaking evidence presented in the "redundant evidence" condition, showing that the strengths of these evidence types did not account for the difference in revision rate between conditions (see SM).

Furthermore, the rational choice model outperformed both the cue saliency and null models. A rational agent in this task would attribute a higher weight to the second instance of new evidence than the second instance of redundant evidence; the parameters for the rational choice model confirmed that this was the case (Fig. 5C and SM). These results confirm that chimpanzees can both remember the specific evidence used to form their beliefs and distinguish between new and redundant evidence when revising their beliefs.

In Experiment 5, we presented chimpanzees ($N = 22$) with second-order evidence (*33*, *34*): evidence that reduces the strength of the first-order evidence (so-called "undercutting defeaters"). This experiment consisted of two conditions: a "defeater" condition (in which first-order evidence was weakened) and a "non-defeater" condition (in which first-order evidence was not weakened). We ran both visual (Fig. 5) and auditory versions (fig. S6) of this task, using the strong evidence from Experiments 1 and 2, respectively. In the "visual defeater" condition, after chimpanzees had chosen the location supported by visual evidence (where they had seen the food), the experimenter pulled a glass

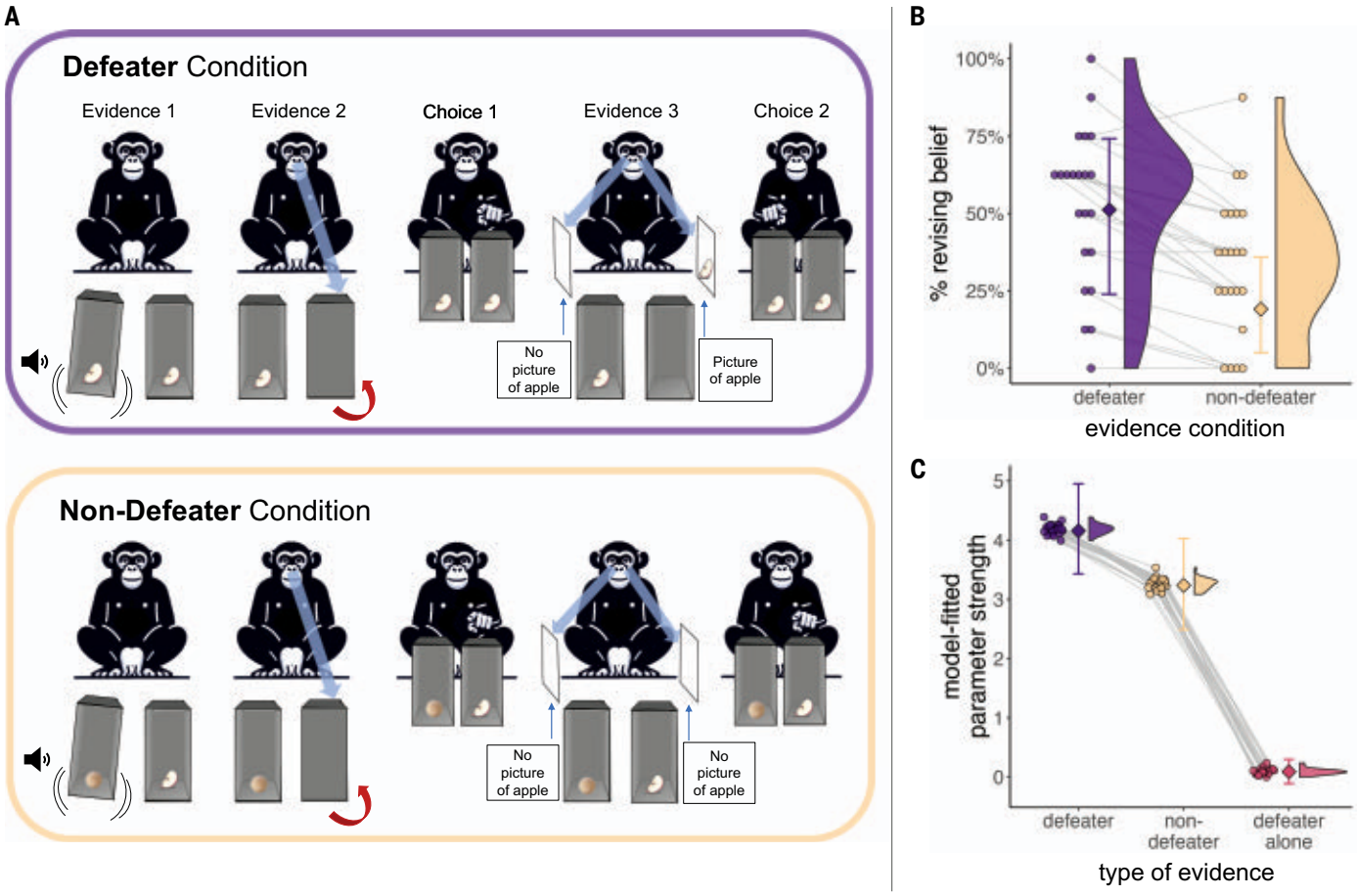


Fig. 5. Procedure and results of Experiment 5. (**A**) In the defeater condition, chimpanzees received weak evidence for one of the boxes, followed by strong evidence for the other box. They then made a first choice, after which they were presented with a defeater: second-order evidence that weakens the strength of the first-order evidence. The strong visual evidence was defeated when the experimenter revealed that the food they had previously seen inside the box was actually a picture on the box's window. In the non-defeater condition, the experimenter's actions were identical, but this time there was no picture on the window. (**B**) Chimpanzees were more likely to revise their belief in the defeater condition than in the non-defeater condition. The figure depicts the combined results of both the visual and auditory versions of the task (see SM for separate figures). (**C**) The rational choice model confirmed that the defeater affected the chimpanzees' beliefs more than the non-defeater, and that the defeater did not affect the chimpanzees' beliefs at all when it was presented alone (without first-order evidence). Note: For clarity, Fig. 5A only depicts the visual version of the task; for the auditory version, please refer to the SM and fig. S6.

window out of the box to reveal that it had a picture of an apple on it (figs. S6, S7, and SM). In the "visual non-defeater" condition, the experimenter also revealed the box's window, which was an empty pane of glass. In the "auditory defeater" condition, after chimpanzees had chosen the location supported by auditory evidence (where they had heard a sound inside the shaking box), the experimenter removed a stone from the box and showed it to the subject (fig. S7 and SM). In the "auditory non-defeater" condition, the experimenter removed a leaf from the box (which could not have caused the rattling sound). We asked whether chimpanzees would be more likely to switch to the box supported by weak evidence in the "defeater" condition, where the strong evidence had been weakened, compared with the "non-defeater" condition, where the strong evidence had not been weakened. The results confirmed this prediction: Chimpanzees were significantly more likely to switch in the "defeater" condition compared with the "non-defeater" condition ($\chi^2 = 16.27$, df = 1, $p < 0.001$). This effect held true in both the auditory (fig. S6d) and visual (fig. S6c) modalities; the interaction of task version and condition was not significant ($\chi^2 = 0.11$, df = 1, $p = 0.743$).

The rational choice model once again strongly outperformed both the cue saliency and null models on this task. Moreover, the best-fitting parameters for the rational choice model confirmed that subjects' beliefs were more affected by the defeater than the non-defeater. Notably, chimpanzees revised beliefs only when defeaters interacted with prior evidence, not when they were presented alone (Fig. 5C and SM).

Taken together, these findings suggest that chimpanzees appropriately respond to second-order evidence—evidence about evidence—which requires representing the evidential relation between evidence and belief.

Discussion

Chimpanzees' responses to counterevidence aligned with a Bayesian model of rational belief revision. They selectively maintained or revised their initial beliefs by evaluating the strength of the initial evidence relative to the strength of the counterevidence. In doing so, they demonstrated remarkable cognitive control and flexibility. Chimpanzees did not attribute a fixed value to each type of evidence; instead, they weighed the relative strength of evidence. For instance,

auditory evidence was judged to be weaker than direct visual evidence in Experiment 1 and stronger than indirect visual evidence in Experiment 2. These responses were enabled by chimpanzees' capacity to represent (i) both possible hypotheses (Experiment 3) and (ii) the respective evidence in favor of them (Experiment 4). Chimpanzees also revised their beliefs in light of undercutting defeaters: second-order evidence that weakens the evidential connection between a piece of first-order evidence and a hypothesis (Experiment 5). Model comparisons demonstrate that these results are not explained in terms of nonrational processes such as belief perseveration, recency bias, or cue saliency. Instead, they suggest that chimpanzees engage in rational belief revision.

Selective belief revision is a metacognitively demanding and reflective cognitive process (*35*). Previous research has shown that chimpanzees appropriately respond to different types of direct and indirect evidence (*6*, *8*), but such evidence-based belief formation can occur in the absence of explicit reasoning about evidence and its bearing on different hypotheses. Indeed, it is often explained in terms of "unreflective evidence-responsiveness" (*33*). By contrast, the capacity to appropriately respond to counterevidence and, especially, to second-order evidence—evidence about the relevance of first-order evidence to one's belief—is thought to constitute a key standard of reflective engagement with evidence (*33*, *35*). It requires representing the three components of the so-called evidential relation: the evidence, the hypothesis, and their causal connection.

These results inform ongoing discussions about the metacognitive capacities of nonhuman animals. It has long been argued that linguistic representation is necessary for successful reflection on evidential relations (*36*–*39*). Yet recent empirical research on information search (*40*, *41*) and opt-out behaviors (*42*–*44*) has led theorists to attribute some metacognitive skills to primates (*45*). Despite this evidence, critics have maintained that behaviors such as searching for information and opting out are explainable in terms of cognition about the world and do not require metacognition—cognition about cognition (*46*).

Our findings present strong support for the view that chimpanzees have genuine metacognitive capacities. The particular form of metacognition demonstrated in the current experiments is "source monitoring": representing not only what one knows but also how one knows it (*47*–*50*). Evidence for this capacity comes from chimpanzees' responses to counterevidence, to new versus redundant evidence, and to second-order evidence. Reasoning about second-order evidence is, by definition, a form of metacognition. That chimpanzees represented second-order evidence as such is most strongly supported by the finding that the undercutting defeater in Experiment 5 affected chimpanzees' decisions only when it followed a piece of first-order evidence but not when presented on its own.

Accounts of major transitions in cognitive evolution identify key changes in the computational architecture of nervous systems that give rise to new cognitive capacities (*45*, *51*, *52*). The emergence of reflective processes—a system able to form representations of the cognitive processes themselves (*51*, *53*)—marks a key transition. The experiments presented here suggest that chimpanzees possess this capacity. Chimpanzees revise their beliefs selectively based on metacognitive awareness of the relation between evidence and hypothesis.

REFERENCES AND NOTES

1. J. Broome, *Rationality Through Reasoning* (Wiley Blackwell, 2013).
2. H. Kornblith, *Scientific Epistemology: An Introduction* (Oxford Univ. Press, 2021).
3. P. Gärdenfors, Ed., *Belief Revision* (Cambridge Univ. Press, ed. 1, 1992).
4. L. A. Heimbauer, R. L. Antworth, M. J. Owren, *Anim. Cogn.* **15**, 45–55 (2012).
5. O. Petit *et al.*, *Anim. Cogn.* **18**, 821–830 (2015).
6. R. Menzel, J. Fischer, Eds., *Animal Thinking: Contemporary Issues in Comparative Cognition*, Strüngmann Forum Reports (MIT Press, 2012).
7. V. Schmitt, J. Fischer, *J. Comp. Psychol.* **123**, 316–325 (2009).
8. C. J. Völter, J. Call, in *APA Handbook of Comparative Psychology: Perception, Learning, and Cognition*, J. Call, G. M. Burghardt, I. M. Pepperberg, C. T. Snowdon, T. Zentall, Eds. (American Psychological Association, 2017), pp. 643–671.
9. M. Tomasello, J. Call, *Primate Cognition* (Oxford Univ. Press, 1997).
10. B. L. Schwartz, M. J. Beran, Eds., *Primate Cognitive Studies* (Cambridge Univ. Press, 2022).
11. K. Andrews, *The Animal Mind: An Introduction to the Philosophy of Animal Cognition* (Routledge, ed. 2, 2020).
12. M. Bohn *et al.*, *Nat. Ecol. Evol.* **7**, 927–938 (2023).
13. C. J. Völter, J. Call, *J. Comp. Psychol.* **128**, 199–208 (2014).
14. Z. Civelek, C. J. Völter, A. M. Seed, *Proc. Biol. Sci.* **288**, 20211101 (2021).
15. D. Hanus, J. Call, *Curr. Biol.* **18**, R370–R372 (2008).
16. J. Call, *J. Comp. Psychol.* **118**, 232–241 (2004).
17. J. M. Engelmann *et al.*, *Child Dev.* **94**, 1102–1116 (2023).
18. K. Karg, M. Schmelz, J. Call, M. Tomasello, *J. Comp. Psychol.* **128**, 431–439 (2014).
19. C. O'Madagain *et al.*, *Proc. Biol. Sci.* **289**, 20212686 (2022).
20. G. Gardiner, *Episteme* **11**, 83–95 (2014).
21. B. Carpenter *et al.*, *J. Stat. Softw.* **76**, 1–32 (2017).
22. C. A. Anderson, *J. Exp. Soc. Psychol.* **19**, 93–108 (1983).
23. C. A. Anderson, K. L. Kellam, *Pers. Soc. Psychol. Bull.* **18**, 555–565 (1992).
24. S. Nestler, *Soc. Psychol.* **41**, 35–41 (2010).
25. F. Chiesi, C. Primi, Recency Effects in Primary-Age Children and College Students, *Int. Elect. J. Math. Ed.* **4**, 259–279 (2009).
26. D. Fudenberg, D. K. Levine, *Proc. Natl. Acad. Sci. U.S.A.* **111** (Suppl 3), 10826–10829 (2014).
27. J. N. Haney, J. L. Bruning, *Psychon. Sci.* **20**, 85–86 (1970).
28. A. M. Treisman, G. Gelade, *Cogn. Psychol.* **12**, 97–136 (1980).
29. M. J. Beran, *J. Comp. Psychol.* **115**, 181–191 (2001).
30. J. Call, *J. Comp. Psychol.* **114**, 136–147 (2000).
31. J. M. Engelmann *et al.*, *Curr. Biol.* **31**, R1377–R1378 (2021).
32. J. M. Engelmann *et al.*, *Biol. Lett.* **19**, 20230179 (2023).
33. G. Melis, S. Monsó, *Philos. Q.* **74**, 844–864 (2023).
34. H. Schleihauf, E. Herrmann, J. Fischer, J. M. Engelmann, *Child Dev.* **93**, 1072–1089 (2022).
35. D. Kuhn, *Educ. Psychol.* **57**, 73–86 (2022).
36. H. Mercier, D. Sperber, *Behav. Brain Sci.* **34**, 57–74 (2011).
37. C. M. Korsgaard, *Fellow Creatures: Our Obligations to the Other Animals*, *Uehiro Series in Practical Ethics* (Oxford Univ. Press, 2018).
38. R. Brandom, *Making It Explicit: Reasoning, Representing, and Discursive Commitment* (Harvard Univ. Press, ed. 4, 2001).
39. T. Burge, C. Peacocke, *Proc. Aristot. Soc.* **96**, 91–116 (1996).
40. J. Call, M. Carpenter, *Anim. Cogn.* **3**, 207–220 (2001).
41. A. G. Rosati, E. Felsche, M. F. Cole, R. Atencia, J. Rukundo, *Cognition* **251**, 105898 (2024).
42. J. D. Smith, J. J. Couchman, M. J. Beran, *J. Comp. Psychol.* **128**, 115–131 (2014).
43. R. R. Hampton, *Anim. Behav. Cogn.* **6**, 230–235 (2019).
44. R. R. Hampton, *Comp. Cogn. Behav. Rev.* **4**, 17–28 (2009).
45. M. Tomasello, *The Evolution of Agency: Behavioral Organization from Lizards to Humans* (MIT Press, 2022).
46. J. Perner, in *Foundations of Metacognition*, M. J. Beran, J. Brandl, J. Perner, J. Proust, Eds. (Oxford Univ. Press, 2012), pp. 94–116.
47. M. K. Johnson, S. Hashtroudi, D. S. Lindsay, *Psychol. Bull.* **114**, 3–28 (1993).
48. A. Gopnik, P. Graf, *Child Dev.* **59**, 1366–1371 (1988).
49. J. Nagel, *Episteme* **12**, 297–308 (2015).
50. L. Cosmides, J. Tooby, in *Metarepresentations*, D. Sperber, Ed. (Oxford Univ. Press, 2000), pp. 53–116.
51. A. B. Barron, M. Halina, C. Klein, *Proc. Biol. Sci.* **290**, 20230671 (2023).
52. P. Godfrey-Smith, *Complexity and the Function of Mind in Nature* (Cambridge Univ. Press, 1996).
53. P. Godfrey-Smith, in *The Philosophy of Daniel Dennett* (Oxford Univ. Press, 2018), pp. 224–253.
54. H. Schleihauf, E. Sanford, Chimpanzees rationally revise their beliefs, Dryad (2025); https://doi.org/10.5061/dryad.3bk3j9kz9.

ACKNOWLEDGMENTS

We thank T. Mukungu, I. Ampeire, P. Sekulya, A. Okello, B. Apamaku, D.y Basemera, H. Aligumisiriza, I. Ayebazibwe, and all the staff at Ngamba Island Chimpanzee Sanctuary. Thanks also go to M. Goddu, A.-C. Schneider, S. Thomaes, and M. Tomasello, who provided helpful feedback on the manuscript. **Funding:** J.M.E. and E.M.S. were supported by a grant from the Leakey Foundation, and H.S. was supported by an Audacity Grant from the Leibniz ScienceCampus Primate Cognition. **Author contributions:** Conceptualization: H.S., J.C., E.H., J.M.E., S.Z., E.M.S.; Investigation: H.S., J.M.E., E.H.; Methodology: H.S., E.M.S., B.D.T., J.C., E.H., J.M.E., S.Z.; Project administration: J.R., E.H., J.M.E.; Resources: J.R., H.S., J.M.E.; Visualization: H.S., E.M.S.; Writing – original draft: J.M.E., H.S., E.M.S.; Writing – review & editing: H.S., E.M.S., B.D.T., J.C., E.H., S.Z., J.M.E. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data and code used in the analyses are available on Dryad (*54*). **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adq5229
Participant Information; Materials and Methods; Supplementary Text; Figs. S1 to S10; References (*55*–*66*); Tables S1 to S20; Movie S1 to S5; MDAR Reproducibility Checklist

Submitted 17 May 2024; resubmitted 22 November 2024; accepted 13 August 2025
10.1126/science.adq5229

EVOLUTION

Genomic architecture of egg mimicry and its consequences for speciation in parasitic cuckoos

Justin Merondun[1,2,3], Frode Fossøy[4], Swetlana Meshcheryagina[5], Phil Atkinson[6], Gennadiy Bachurin[7], Victor Bulyuk[8]†, Viktar Fenchuk[9], Mikhail Golovatin[5], Chris Hewson[6], Marcel Honza[10], Mikhail Markovets[8], Csaba Moskát[11], Gregory L. Owens[3], Petr Procházka[10], Yaroslav Red'kin[12,13], Jarkko Rutila[14], Michal Šulc[10], Kasper Thorup[15], Bård G. Stokke[16], Wei Liang[17], Jochen B. W. Wolf[1,2]*

Host-parasite arms races facilitate rapid evolution and can fuel speciation. *Cuculus* cuckoos are deceptive egg mimics that exhibit a broad diversity of counterfeit egg phenotypes, representing host-adapted subpopulations (gentes). Genome analysis of 298 common (*Cuculus canorus*) and 50 oriental cuckoos (*Cuculus optatus*) spanning 15 egg morphs revealed that eggshell background coloration is predominantly influenced by matrilineal genetic variation. Recurrent mitochondrial mutations and an ancient W chromosome–linked translocation of an autosomal assembly factor for respiratory complex I provide a tentative link between mitochondrial function and pigment synthesis through the heme pathway. Biparentally inherited loci contribute to phenotypic variation in both species, mainly for maculation. The evolutionary tug-of-war over a sex-limited, mimetic trait integrates autosomal components with the nonrecombining, matrilineal genome without catalyzing genome-wide divergence between gentes.

The common cuckoo (*Cuculus canorus*) is a textbook example of an obligate brood parasite that has fascinated naturalists since the birth of zoology (*1*, *2*). It exploits more than 100 host species (*3*, *4*) and displays large variation in mimetic egg phenotypes. It has been suggested that host-matched phenotypes evolved in subpopulations specialized in parasitizing a particular host species while maintaining evolutionary independence (host races or gentes) (*5*, *6*). Gens diversification is believed to be fueled by concurrent coevolutionary arms races between cuckoos and their passerine hosts, many of which are capable of rejecting parasitic eggs that are dissimilar from their own (*3*, *7*). Yet for the parasite, the expected evolutionary outcome strongly depends on the genetic architecture and inheritance of the traits involved (*8*, *9*).

Host switches can lead to reproductive isolation and promote speciation if the mimetic trait encompasses both sexes and is encoded on the autosomal genome and if mating is assortative by host (*10*, *11*). This is exemplified in indigobirds, whydahs, and bronze cuckoos, where host mimicry involves nestling traits expressed by both sexes (*12*, *13*). In combination with host-specific assortative mating mediated by imprinting [e.g., on host song (*14*)], coevolution has facilitated rapid diversification of parasitic lineages (*13*, *15*). Alternatively, female-limited parasitic traits, if encoded on the matrilineal component of the genome (mtDNA, W chromosome), can be stably inherited from mothers to daughters irrespective of assortative mating (*6*, *16*). Such a genetic architecture is less conducive to speciation (*8*) and has been suggested for honeyguides (*17*), cuckoo finches (*18*), and in small-scale genetic studies for common cuckoos (*3*, *19*–*22*).

Here, we revisit the two competing hypotheses governing diversification of host races in two brood parasitic *Cuculus* species: the common cuckoo (*C. canorus*; *canorus* hereafter) and the closely related oriental cuckoo (*C. optatus*; *optatus* hereafter) (*23*). We assess in both species whether host races, operationally defined by egg phenotype, are predicated on matrilineal variation (*3*, *19*–*22*) or whether biparental inheritance contributes to egg diversity (*24*, *25*).

Geographic structure of genetic variation

We characterized both matrilineal (mtDNA, W chromosome) and coinherited (autosomes, Z chromosome) genetic variation across the distributional range of 202 unrelated *canorus* and 50 *optatus* representing a total of 25 host species (21 *canorus*, 4 *optatus*) and 15 egg morphs (11 *canorus*, 4 *optatus*) (fig. S1 and table S1). Matrilineal phylogenies indicate seven primary haplogroups in *canorus* ($M_{CC}1$ to $M_{CC}7$) and four in *optatus* ($M_{CO}1$ to $M_{CO}4$) (Fig. 1A and fig. S2). An ancient split 1.4 million to 2.9 million years ago, exceeding *optatus* species boundaries [<1 million years (*23*)], separates east Asian *canorus* from European *canorus* (minimum and maximum 95% highest posterior density; figs. S3 to S5). An exception to this East-West geographic divide is the *canorus* lineage of blue egg–laying cuckoos in Finland ($M_{CC}1$) clustering with haplogroups $M_{CC}2$ and $M_{CC}3$ predominantly found in East Asia (*19*). Among the remaining four haplogroups ($M_{CC}4$ to $M_{CC}7$) of the *canorus* lineage, $M_{CC}4$ is closely related to *optatus* lineages $M_{CO}1$ and $M_{CO}2$. These haplogroups correspond exclusively to females of the rufous plumage morph, which, similar to egg phenotypic variation, appears to be subject to balancing selection (*23*) (Fig. 1A and figs. S2 and S3). In *canorus*, haplogroups are broadly spread across the species range, except for East Asian $M_{CC}3$ and the blue egg–associated $M_{CC}1$ of Finnish *canorus* (Fig. 1C). In *optatus*, $M_{CO}1$ is maintained across the distributional range, whereas the other haplogroups are geographically restricted (Fig. 1D and fig. S6).

Autosomal population structure in *canorus* comprises five major genetic clusters [$A_{CC}1$ to $A_{CC}5$, maximum fixation index (F_{ST}) = 0.085], each aligning with broad geographic areas across Eurasia (Fig. 1B and figs. S7 and S8). At finer resolution among 14 geographically defined groups (geographic clusters $G_{CC}1$ to $G_{CC}14$; fig. S9), significant isolation by distance is evident (Spearman's $\rho = 0.92$, $p = 2.4 \times 10^{-18}$). By contrast, matrilineal variation shows much weaker isolation by distance (W chromosome $\rho = 0.33$, $p = 0.04$) (Fig. 1, C and E, and fig. S10), counter to expectations from reduced effective population size but consistent with female-biased dispersal (*26*, *27*). These results highlight an association primarily between geographic location and biparental autosomal variation (Fig. 1E). Meanwhile, deeply divergent haplogroups are widely dispersed across the *canorus* range, offering an opportunity for host-specific selection to exploit segregating matrilineal variation (Fig. 1C and fig. S11). This pattern is reversed within *optatus*, where autosomal variation is less associated

[1]Department of Evolutionary Biology, LMU Munich, Planegg-Martinsried, Germany. [2]Department of Microevolution and Biodiversity, Max Planck Institute for Biological Intelligence, Seewiesen, Germany. [3]Department of Biology, University of Victoria, Victoria, BC, Canada. [4]Centre for Biodiversity Genetics, Norwegian Institute for Nature Research, Trondheim, Norway. [5]Institute of Plant and Animal Ecology, Ural Branch, Russian Academy of Sciences, Yekaterinburg, Russia. [6]British Trust for Ornithology, Thetford, UK. [7]Scientific and Practical Center of Biodiversity, Irbit, Russia. [8]Rybachy Biological Station, Zoological Institute, Russian Academy of Sciences, Rybachy, Russia. [9]The Frankfurt Zoological Society, Frankfurt, Germany. [10]Institute of Vertebrate Biology, Czech Academy of Sciences, Brno, Czech Republic. [11]Hungarian Natural History Museum, Budapest, Hungary. [12]Zoological Museum, Moscow State University, Moscow, Russia. [13]Institute of Biological Problems of the Cryolithozone, Siberian Branch, Russian Academy of Sciences, Yakutsk, Russia. [14]Department of Environmental and Biological Sciences, University of Eastern Finland, Joensuu, Finland. [15]Center for Macroecology, Evolution and Climate, Globe Institute, University of Copenhagen, Copenhagen, Denmark. [16]Department of Terrestrial Ecology, Norwegian Institute for Nature Research, Trondheim, Norway. [17]Ministry of Education Key Laboratory for Ecology of Tropical Islands, Hainan Normal University, Haikou, People's Republic of China. ***Corresponding author. Email: j.wolf@bio.lmu.de** †Deceased

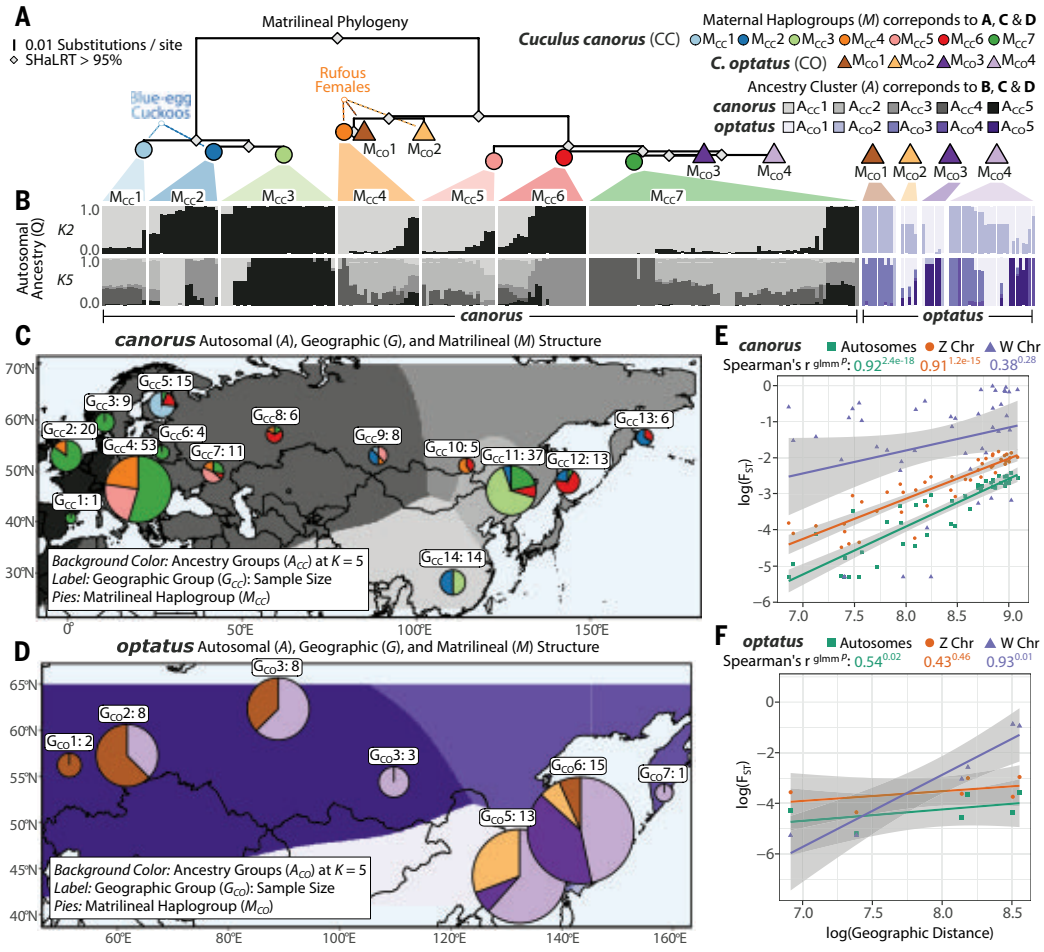


Fig. 1. Genetic structure across the geographic range of *C. canorus* and *C. optatus*. (**A**) Matrilineal phylogeny of the W chromosome (N = 51,605 SNPs) with truncated branches identifying seven *canorus* (CC) maternal haplogroups ($M_{CC}1$ to $M_{CC}7$) and four *optatus* (CO) haplogroups ($M_{CO}1$ to $M_{CO}4$). Diamonds indicate nodes with Shimodaira-Hasegawa approximate likelihood ratio test (SHaLRT) support >95%. (**B**) Ancestry analysis (n = 202 *canorus*, 50 *optatus*) inferred from unlinked autosomal SNPs (N = 1,546,013 *canorus*, 920,854 *optatus*), arranged by maternal haplogroup. (**C**) Spatial structuring of *canorus*. Background color (grayscale; fig. S6) indicates the spatial distribution of autosomal ancestry ($A_{CC}1$ to $A_{CC}5$) inferred using ADMIXTURE at (K) = 5. Pie charts indicate maternal haplogroups, with size corresponding to sample size. Samples are clustered into 14 geographic groups ($G_{CC}1$ to $G_{CC}14$) by k-means on latitude and longitude; group labels and sizes are shown above the pie charts. Please see supplementary materials for alternative symbology (figs. S8 and S11). (**D**) Spatial structuring of *optatus*. (**E**) Correlations between geographic distance (km) and genetic distance in *canorus*, summarized from autosomal (F_{ST}) or matrilineal (Φ_{ST}) variation across pairwise comparisons between geographic groups. Spearman's rank correlation (r_s) and generalized linear mixed model (glmm) Bonferroni-adjusted p (shown in superscript) for each genomic compartment are indicated above, with shading indicating 95% confidence intervals. (**F**) Correlations for *optatus*.

with geographic distance (Spearman's $\rho = 0.54$, $p = 1.0$) than matrilineal variation (W chromosome $\rho = 0.93$, $p = 0.03$) (Fig. 1, D and F, and fig. S10).

Genomic correlates of egg diversity

Next, limiting the dataset to females with known maternal egg phenotype (E_{CC}: n = 60 *canorus*; E_{CO}: n = 27 *optatus*), we assessed whether *canorus* and *optatus* egg diversification follows strict matrilineal inheritance (*5*, *19*, *21*, *28*) (M_{CC}, M_{CO}) or is better explained by autosomal contributions (*24*, *25*) (A_{CC}, A_{CO}) or geographic location reflecting local host availability (G_{CC}, G_{CO}) (Fig. 2, A and B, and fig. S12). We replicated the following analyses with an expanded dataset that included males (n = 86 *canorus*, 50 *optatus*), which yielded qualitatively identical patterns despite restricting the matrilineal component to mitochondria (figs. S13 and S14).

Within *optatus*, where egg phenotype variation is limited to differences in maculation on shared eggshell coloration and egg size (*29*), geographic location is the primary predictor (distance-based redundancy analysis pseudo-adjusted R^2 = 0.78), with weaker associations for matrilineal structure (R^2 = 0.04) and autosomal variation (R^2 = 0.02) (Fig. 2C). This finding is supported by classical machine learning, as adding a matrilineal component to geographic and autosomal predictors did not improve model predictions (Fig. 2F and fig. S13). By contrast, variation in *canorus* egg morphs, differing in both eggshell coloration and maculation, are most strongly associated with matrilineal variation (distance-based redundancy analysis pseudo-adjusted R^2 = 0.26), followed by geographic location (R^2 = 0.16) and autosomal ancestry (R^2 = 0.02) (Fig. 2, A and C, and fig. S14). Matrilineal inheritance is further supported by kinship analyses of 107 related *canorus* individuals: Phenotype concordance is complete among maternal relatives (100% across all kinship levels), whereas paternal relatives, particularly diagnostic cousins, show significantly less concordance (75.8%; n = 50/66 matches; Fisher's exact test p = 0.01) (Fig. 2D and fig. S15). Machine learning results reinforce the primacy of matrilineal inheritance, as adding haplogroup information substantially improves classification of *canorus* egg morphs (Fig. 2E). The effect persisted, though diminished,

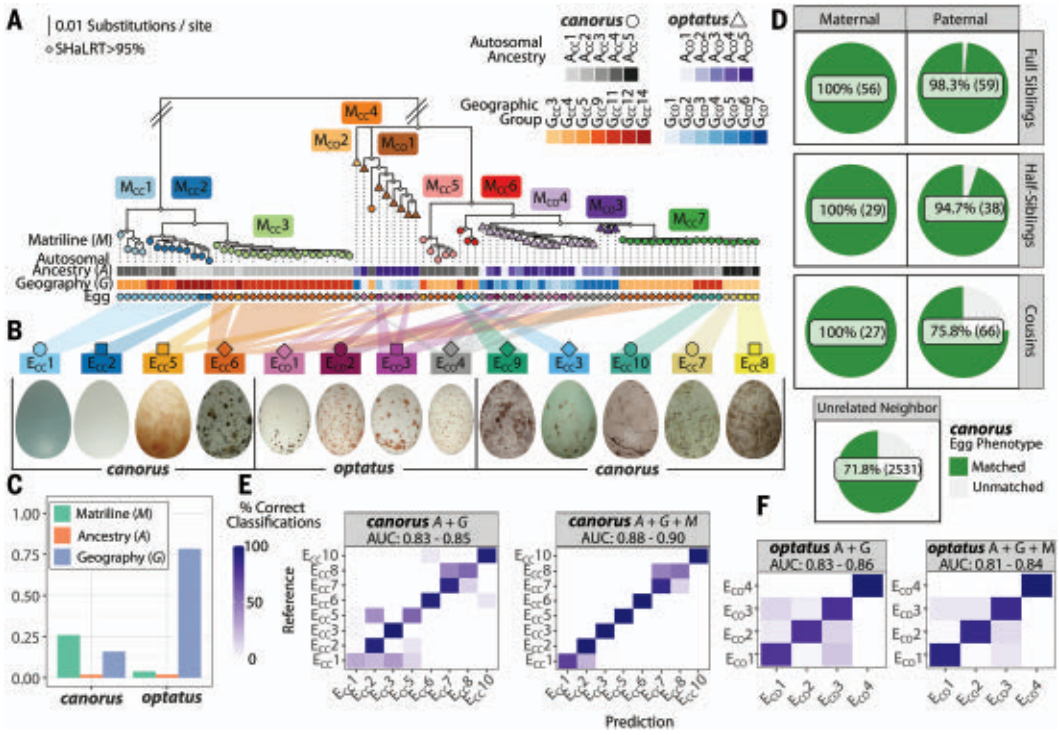


Fig. 2. Matrilineal variation is associated with egg diversification in *C. canorus*. (**A**) Matrilineal W chromosome phylogeny ($n = 60$ *canorus*, 27 *optatus*), with corresponding autosomal groups (*A*), geographic groups (*G*), and egg morphs indicated below each tip, each representing an unrelated female with known egg morph. Diamonds indicate node support exceeding 95% (SHaLRT). (**B**) Phenotypic variation within *canorus* (E_{CC}) and *optatus* (E_{CO}) egg morphs. (**C**) Variation explained (pseudo-R^2) by continuous pairwise distance matrices of matrilineal, geographic, or autosomal variation estimated with a distance-based redundancy analysis (compare figs. S13 and S14 for replicated analyses with mtDNA variation). (**D**) Egg phenotype matching among varying degrees of related *canorus*. (**E**) Egg morph predicted with machine learning using either autosomal ancestry cluster (*A*) and geographic group (*G*) alone (left) or with the addition of haplogroup (*M*) (right). Confusion matrix of predictive modeling shows the proportion of correct classifications across 100 bootstrapped replicates. The area under the curve (AUC) 95% confidence interval across replicates is indicated below each label. (**F**) Confusion matrix for *optatus*.

when divergent $M_{CC}1$, $M_{CC}2$, and $M_{CC}3$ haplogroups, including the immaculate blue egg ($E_{CC}1$) morph (*19*), were excluded (fig. S14G). However, subtle differences among most brownish egg morphs ($E_{CC}6$, $E_{CC}7$, $E_{CC}8$), which are largely associated with common *Acrocephalus* hosts, remain poorly resolved by any model (Fig. 2, B and E, and figs. S1 and S13). Stochastic character mapping of egg state on maternal and biparental phylogenies provided additional detail. By leveraging mitoautosomal discordance, it corroborated strong matrilineal association for both blue egg types $E_{CC}1$ and $E_{CC}3$ (figs. S16 to S18). Although $E_{CC}1$ has previously been linked to a deeply divergent mitochondrial lineage associated with blue eggs [see (*19*)], greater parsimony of egg-state transitions on the mitochondrial tree for $E_{CC}3$ suggests matrilineal association for a second, distinct blue-green egg phenotype (associated with brambling hosts; Fig. 2E and figs. S16 and S17).

Overall, these findings provide evidence for a matrilineal contribution of egg diversification in *canorus*, likely associated with changes in eggshell coloration evolving from a brownish, $E_{CC}6$-like ancestor (figs. S16 and S17), a widespread morph that allows exploitation of broadly distributed *Acrocephalus* hosts across the species range (*28*, *30*–*32*). Male mate choice appears largely opportunistic and shaped by local host availability (*3*, *5*, *25*, *28*, *33*, *34*), preventing stable genome-wide autosomal transmission through the patriline (*8*, *10*) (Fig. 2, C and E). Egg patterning, which distinguishes many of the nonancient haplogroup morphs within *canorus* and accounts for basically all variation in *optatus* samples, also covaries with geographic structure, enabling adaptation to locally abundant and stable hosts through biparentally inherited genetic variation.

Genetic basis of egg phenotype

Next, we sought to identify candidate genes associated with egg phenotype. Under the assumption that genomic regions with elevated genetic differentiation may flag genetic variation related to egg morph (*35*), we performed an all-versus-one F_{ST} scan with resampling [bootstrapped F_{ST} (bF_{ST}) in 10-kb windows, 100- base pairs for mtDNA]. Consistent with the genome-wide analysis described earlier in the text, matrilineal elements (W chromosome, mtDNA) are enriched for the blue $E_{CC}1$ morph, with W-linked enrichment relative to other egg morphs being more than twofold the permuted expectation (observed/permuted = 2.1, $p < 0.01$) (Fig. 3A and table S2). Conversely, significantly more autosomal outliers than expected were observed for $E_{CC}10$ (enrichment = 1.9; $p < 0.01$). By contrast, *optatus* egg morphs show significant enrichment only on autosomes and the Z chromosome (Fig. 3A).

Genes in close proximity to the main outlier regions include credible functional candidates that influence the two main components of egg phenotype variation: eggshell coloration and maculation. Eggshell color variation, which is more similar between closely related females (*36*), results from the mixing of two heme-derived pigments: protoporphyrin IX (reddish-brown) and biliverdin (blue-green) (*37*). Maculation forms as the egg passes through the uterus and is shaped by the distribution of pigment glands, timing of pigment release, and mode of egg rotation (*37*, *38*).

Of the biparentally inherited genes associated with differences in maculation, several relate to female reproductive function (table S3). In *canorus*, the top differentiated peak for $E_{CC}10$ lies close to *RLN3*, which is involved in pituitary and oviduct signaling in birds (*39*). In

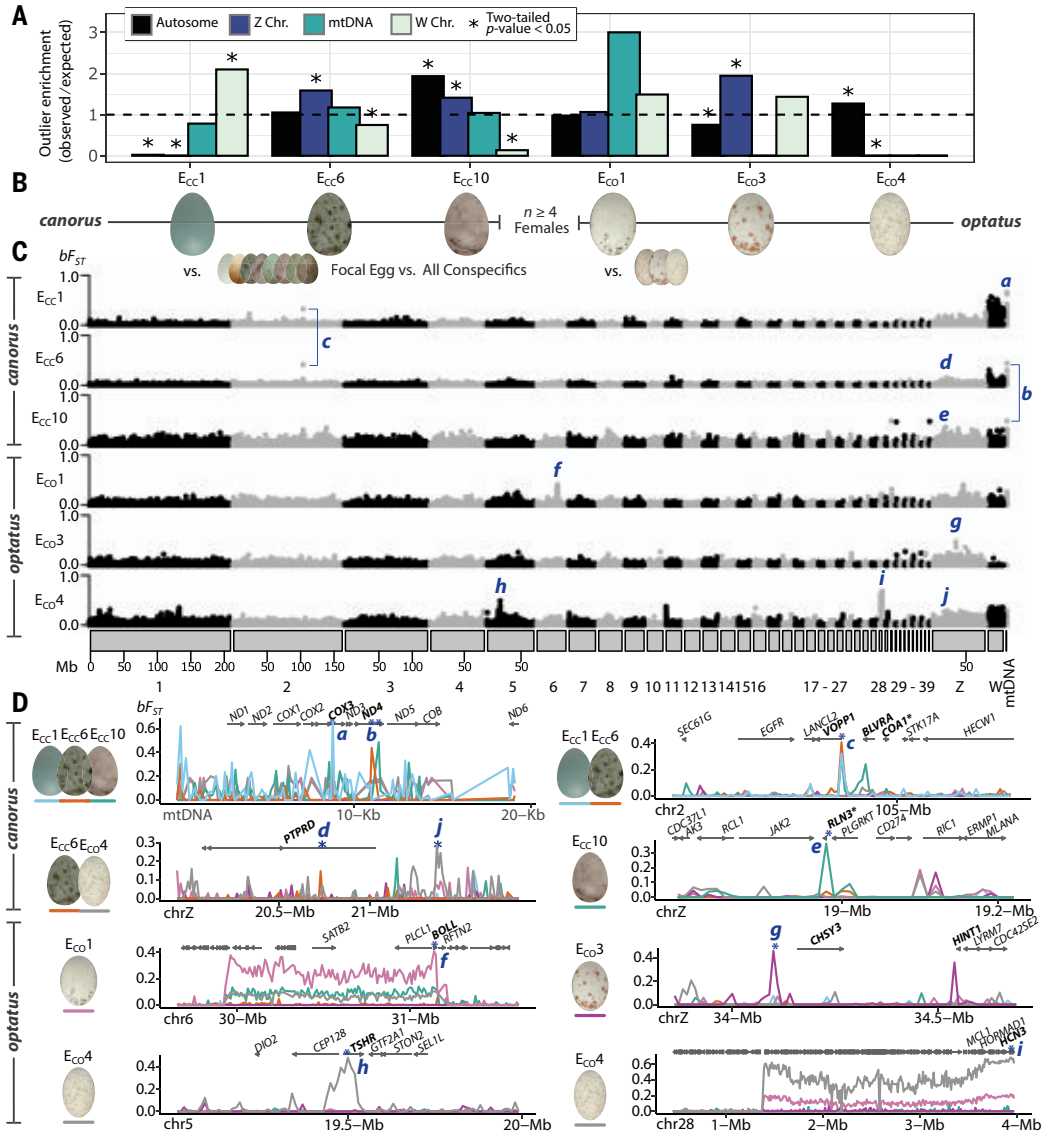


Fig. 3. Genetic basis of egg phenotype. (**A**) Permutation enrichment of bF_{ST} outliers per egg across chromosome class conducted separately for species. For joint analysis across species, see fig. S26 (Note: bF_{ST} is used throughout figure.) (**B**) Pairwise one-versus-all bF_{ST} comparisons, separated by species (*canorus* and *optatus*). (**C**) Genome-wide bF_{ST} in 10-kb windows, with the highest outlier windows indicated with blue letters. (**D**) bF_{ST} peak regions and nearby candidate genes (bold face), with the highest F_{ST} peak indicated with an asterisk. Line colors correspond to different egg morphs, with the primary candidate morph to the left of each plot. Black asterisks indicate unnamed genes with homology inferred from National Center for Biotechnology Information (NCBI) annotations (e.g., LOC104063035 = "relaxin-3"; LOC104065962 = "*COA1*-like"). The blue letters indicate the same outlier windows as in (C).

optatus, key outliers are linked to *BOLL* and *CHSY3* [involved in germ cell development (*40*, *41*)], *HINT1* [linked to embryogenesis and sex hormones (*42*)], and *TSHR* and *HCN3* [involved in hormone signaling and eggshell traits (*43*, *44*)] (table S3). Two regions show recombination suppression, potentially preserving trait combinations (fig. S19). *PTPRD*, linked to eggshell color in domestic chickens (*45*), is differentiated in both $E_{CC}6$ (*canorus*) and $E_{CO}4$ (*optatus*) within ~500 kb of the same locus (Fig. 3D). Secondary peaks across egg morphs also implicate genes in female reproductive development (fig. S20).

In *canorus*, biparentally inherited genes linked to eggshell coloration are either directly involved in the heme biosynthetic pathway or in mitochondrial homeostasis, which regulates this process through metabolic signaling and shared biosynthetic resources (*46*–*48*). Together with studies implicating oxidative stress in female avian reproduction (*49*, *50*), this suggests a tentative, mechanistic link of mitochondrial function and eggshell coloration, which is also observed in chickens (*51*). Consistent with this hypothesis, candidate regions for $E_{CC}1$ and $E_{CC}6$ lie within 170 kb of *BLVRA*, a key gene in heme catabolism that reduces oxidative stress by converting biliverdin to bilirubin (*52*), as well as *VOPP1* and *COA1*, which are both likewise linked to mitochondrial function (*53*, *54*) (Fig. 3, C and D). Analyses replicated using a relaxed threshold requiring three female nestlings for focal contrasts corroborated the association of *BLVRA* and *RLN3* and, for $E_{CC}8$, additionally implicated an autosomal gene involved in mitochondrial fission and associated with egg laying in chickens (*55*) (fig. S21 and table S3).

Isolating the matrilineal mutations for eggshell coloration is challenging owing to the large number of substitutions that have accumulated along long branches and the complete linkage of mitochondrial and W chromosome haplotypes. To identify candidate mutations, we

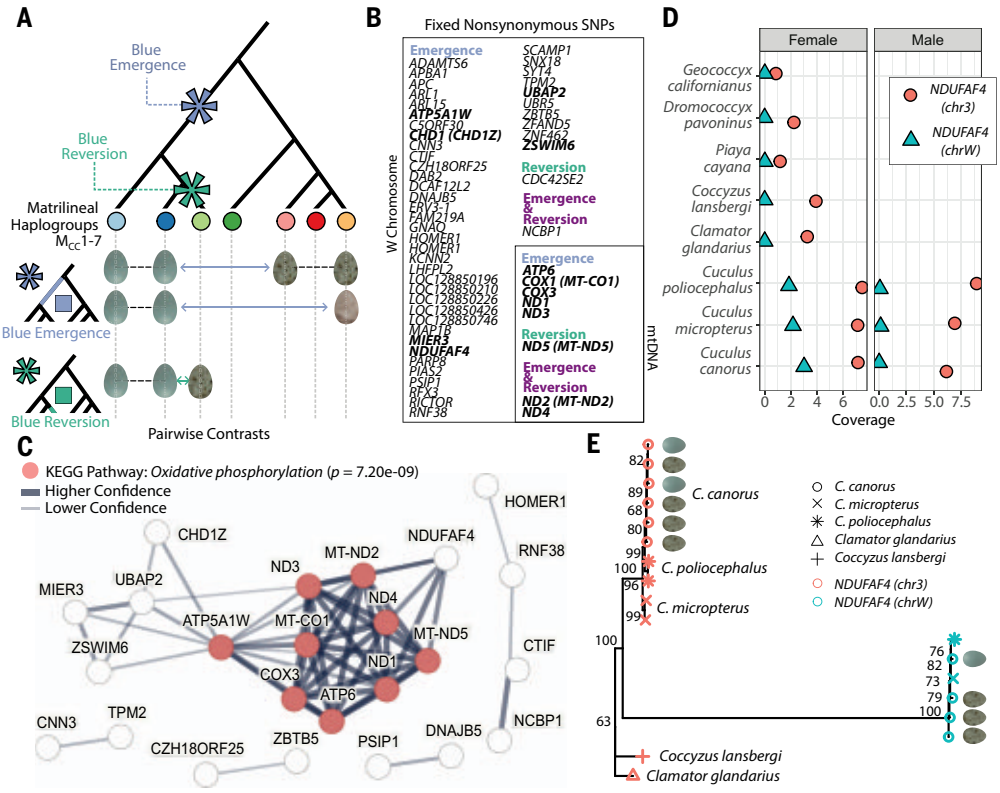


Fig. 4. *C. canorus* blue egg background predominantly maps to oxidative phosphorylation genes. (**A**) Schematic (top) depicting the evolutionary emergence and subsequent diversification of *canorus* egg morphs, with pivotal phenotypic shifts indicated by asterisks. Fixed differences in pairwise comparisons of haplogroups associated with same versus different egg types flag candidate mutations that contribute to phenotypic divergence. Investigated scenarios include the origin of the immaculate blue egg phenotype (blue emergence) and its subsequent reversal (blue reversion). Illustrative schematics (bottom) denote specific branches and mutations within the phylogeny where these evolutionary events are hypothesized to occur, along with the pairwise egg contrasts used to identify them. (**B**) Candidate genes containing a fixed nonsynonymous SNP for each contrast across the W chromosome and mtDNA. Genes in bold are shown in the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway in (C). (**C**) Gene network associations among candidate genes. Note that all nonmitochondrial genes are found on the W chromosome (table S6). (**D**) Resequencing coverage of the autosomal and W chromosome paralog of *NDUFAF4* across Cuculiformes. (**E**) Maximum-likelihood phylogenetic tree of the *NDUFAF4* paralogs, with bootstrap support indicated at each node.

screened for fixed differences between pairs of matrilines associated with an egg type ("matriline association mapping," e.g., $M_{CC}1$:$E_{CC}1$ versus $M_{CC}7$:$E_{CC}6$) (Fig. 4A). We focused on the immaculate blue egg morph ($E_{CC}1$), which occurs across two maternal haplogroups ($M_{CC}1$ and $M_{CC}2$), persists thousands of kilometers apart, and yet remains at high frequencies in the respective populations, putatively owing to highly discriminatory hosts (Fig. 1C) (*3*, *56*). Based on our matrilineal phylogenetic reconstruction, mutations along a long, ancient branch conferred the immaculate blue egg (blue emergence), followed by a reversion within this lineage to other eggshell colors in East Asia (blue reversion) (Fig. 4A). The remaining nonblue haplogroups (e.g., $M_{CC}4$ to $M_{CC}7$) diversified into other egg morphs but never to immaculate blue, based on present sampling and knowledge (Figs. 1A and 2A).

Consistent with the broad-stroke outlier analyses (Fig. 3), fixed single-nucleotide polymorphisms (SNPs) are predominantly observed in the matrilineal genome across both evolutionary scenarios. This is particularly pronounced for the origin of the blue egg (blue emergence), where 55% (n = 123/225) of mtDNA and 75% (n = 33,597/44,739) of W chromosomal SNPs are fixed, which is in stark contrast to none on the autosomes or Z chromosome (table S4). This pattern is reiterated for the reverted $E_{CC}6$ egg contrast (blue reversion) with 12% (n = 9/73) fixed SNPs in mtDNA, 5% (n = 413/8,490) in the W chromosome, and none on the autosomes or Z chromosome. Mitochondrial genes *ND2* and *ND4* harbor nonsynonymous SNPs linked to both emergence and reversion of blue eggshell coloration (Fig. 4B), a pattern corroborated with de novo mitochondrial assemblies that further implicates *COX1* and *ND5* (figs. S22 and S23). Together with biparentally inherited genes associated with mitochondrial homeostasis (Fig. 3D), these fixed nonsynonymous mutations could be implicated in mitochondrial regulation for eggshell coloration.

Although a mitochondrial shift in egg phenotype mediated by the heme pathway is mechanistically plausible, several W-linked genes containing a fixed nonsynonymous SNP also interact with the statistically enriched oxidative phosphorylation pathway ($p = 7.20 \times 10^{-9}$) (Fig. 4C and table S5). Among them, an *NDUFAF4*-like gene, a homolog of *NDUFAF4* acting in the respiratory complex I assembly, represents a candidate for further functional investigation (LOC128850245 in RefSeq Assembly GCF_017976375.1; table S6). It has a copy on the W chromosome (28 kb; mean female coverage = 6.4×, male coverage = 0.13×) in addition to an ancestral autosomal copy on chromosome 3 (4.0 kb; female coverage = 11×, male = 11×), which exhibits no autosomal differentiation among *canorus* egg morphs (Fig. 4D and fig. S24). Phylogenetic analysis across the Cuculiformes order suggests that the W paralog was acquired at the onset of the brood parasitic genus *Cuculus* (Fig. 4E and table S6). The W-linked copy contains the same number of exons and has full open reading frames (fig. S25) yet is substantially diverged from its autosomal paralog since insertion in the *Cuculus* ancestor (Fig. 4E). Although mitonuclear cross-talk has broad biological implications (*57*), our results suggest the possibility that polygenic autosomal–mitochondrial–W chromosome interactions contribute to egg background color diversification. Other avian species have co-opted autosomal genes to produce immaculate blue eggs, including SNPs upstream of efflux transporter *ABCG2* in ducks (*58*), an endogenous retroviral insertion upstream of anion transporter *SLCO1B3* in chickens (*59*), and a recessive gene in quail (*60*), illustrating the multiple mutational pathways that underlie egg diversification in birds.

Conclusions

In *canorus*, phenotypic variation in egg traits, especially eggshell coloration, appears primarily mediated by matrilineally inherited genes, as suggested nearly a century ago (*6*). This remains surprising, as exclusive matrilineal inheritance poses great challenges: Reduced effective population size limits the efficacy of selection, and beneficial mutations compete among matrilines rather than recombining into advantageous trait combinations. Thus, barring new mutations, we would not expect, for example, a female with rufous plumage (associated with

haplogroup $M_{CC}4$) to lay immaculate blue eggs (associated with $M_{CC}1$ and $M_{CC}2$), supporting earlier speculations by Kunz that plumage coloration is linked to host choice and egg coloration (*61*). Additionally, the limited mutational target on mtDNA and the W chromosome may require recruitment of autosomal genes, as suggested here for the translocation of *NDUFAF4*. Despite these constraints, female-biased gene flow involving multiple matrilines maintained by balancing selection (*23*) appears to provide sufficient adaptive variation for exploiting multiple host species.

Biparental genetic contributions and local geographic structure also seem to contribute to phenotypic diversity. In *canorus*, co-option of nuclear genes with mitochondrial-related functions suggests that interactions between matrilineal and autosomal variation may contribute to eggshell coloration. By contrast, in *optatus*, variation is limited to maculation with exclusive contribution of autosomal genes. Such polygenic adaptation—which is also observed in egg phenotype variation of the village weaver (*Ploceus cucullatus*), a host species of Diederik cuckoo that exhibits extensive intra- and interclutch diversity as a defense against parasitism (*62*)—may develop in *canorus* over prolonged coexistence with abundant, accessible host species, such as *Acrocephalus* warblers (*28*, *30*–*32*). Although selective pressures differ for hosts and parasites, the involvement of many unlinked autosomal loci in both systems may reflect convergent selection on egg phenotype diversity.

Host race diversification in the common cuckoo seems largely predicated on matrilineal variation, limiting the potential for speciation. Though biparentally inherited loci contribute to regional egg diversity, mostly for maculation in *optatus*, they fail to establish strong, genome-wide barriers to gene flow that would place gentes on independent evolutionary trajectories.

REFERENCES AND NOTES

1. Aristotle, *Historia Animalium*, W. T. D'Arcy, Transl. (Clarendon Press, 1910).
2. E. Jenner, *Philos. Trans. R. Soc. London* **78**, 219–237 (1788).
3. A. Moksnes, E. Røskaft, *J. Zool.* **236**, 625–648 (1995).
4. R. B. Payne, *The Cuckoos*, vol. 15, Bird Families of the World Series (Oxford Univ. Press, 2005).
5. M. de L. Brooke, N. B. Davies, *Nature* **335**, 630–632 (1988).
6. R. C. Punnett, *Nature* **132**, 892–893 (1933).
7. M. C. Stoddard, M. Stevens, *Evolution* **65**, 2004–2013 (2011).
8. G. Dillenseger, *J. Avian Biol.* **2024**, 03252 (2024).
9. B. Sinervo, E. Svensson, *Heredity* **89**, 329–338 (2002).
10. O. Krüger, M. Kolss, *J. Evol. Biol.* **26**, 2447–2457 (2013).
11. O. Krüger, M. D. Sorenson, N. B. Davies, *Proc. Biol. Sci.* **276**, 3871–3879 (2009).
12. N. E. Langmore *et al.*, *Proc. Biol. Sci.* **278**, 2455–2463 (2011).
13. M. D. Sorenson, K. M. Sefc, R. B. Payne, *Nature* **424**, 928–931 (2003).
14. R. B. Payne, L. L. Payne, J. L. Woods, M. D. Sorenson, *Anim. Behav.* **59**, 69–81 (2000).
15. N. E. Langmore *et al.*, *Science* **384**, 1030–1036 (2024).
16. R. C. Jensen, *Nature* **209**, 827 (1966).
17. C. N. Spottiswoode, K. F. Stryjewski, S. Quader, J. F. R. Colebrook-Robjent, M. D. Sorenson, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17738–17742 (2011).
18. C. N. Spottiswoode *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2121752119 (2022).
19. F. Fossøy *et al.*, *Nat. Commun.* **7**, 10272 (2016).
20. H. L. Gibbs, M. de L. Brooke, N. B. Davies, *Proc. Biol. Sci.* **263**, 89–96 (1996).
21. H. L. Gibbs *et al.*, *Nature* **407**, 183–186 (2000).
22. K. Marchetti, H. Nakamura, H. L. Gibbs, *Science* **282**, 471–472 (1998).
23. J. Merondun *et al.*, *Sci. Adv.* **10**, eadl5255 (2024).
24. F. Fossøy *et al.*, *Proc. Biol. Sci.* **278**, 1639–1645 (2011).
25. J.-W. Lee, H.-K. Moon, H.-J. Noh, M.-S. Kim, J.-C. Yoo, *Behav. Ecol.* **32**, 248–256 (2021).
26. P. J. Greenwood, *Anim. Behav.* **28**, 1140–1162 (1980).
27. R. J. Petit, L. Excoffier, *Trends Ecol. Evol.* **24**, 386–393 (2009).
28. N. Wang *et al.*, *Integr. Zool.* **20**, 144–159 (2025).
29. S. G. Meshcheryagina, A. Mashanova, G. N. Bachurin, I. S. Mitiay, M. G. Golovatin, *J. Zool.* **306**, 147–155 (2018).
30. O. Kleven *et al.*, *J. Avian Biol.* **35**, 394–398 (2004).
31. N. B. Davies, M. de L. Brooke, *Anim. Behav.* **36**, 262–284 (1988).
32. C.-C. Yang *et al.*, *Zool. Res.* **35**, 67–71 (2014).
33. J. Koleček, P. Procházka, V. Brlík, M. Honza, *Naturwissenschaften* **107**, 12 (2020).
34. C. Yang, W. Liang, A. P. Møller, *Behav. Ecol. Sociobiol.* **72**, 126 (2018).
35. R. C. Lewontin, J. Krakauer, *Genetics* **74**, 175–195 (1973).
36. M. Šulc *et al.*, *Zool. J. Linn. Soc.* **195**, 33–44 (2022).
37. G. Y. Kennedy, H. G. Vevers, *Comp. Biochem. Physiol. B* **55**, 117–123 (1976).
38. M. C. Stoddard, *Curr. Biol.* **32**, R1126–R1132 (2022).
39. C. Lv *et al.*, *Front. Physiol.* **13**, 1010851 (2022).
40. R. Rosario, A. J. Childs, R. A. Anderson, *Stem Cell Res.* **21**, 193–201 (2017).
41. C. Lian *et al.*, *In Vitro Cell. Dev. Biol. Anim.* **51**, 690–696 (2015).
42. Y.-P. Lin *et al.*, *Theriogenology* **74**, 1161–1178.e1, 8 (2010).
43. A. Fallahshahroudi *et al.*, *Genetics* **217**, 1–9 (2021).
44. W. Li *et al.*, *Poult. Sci.* **103**, 103715 (2024).
45. R. Liao *et al.*, *Anim. Genet.* **47**, 588–596 (2016).
46. Y. Y. Yien, M. Perfetto, *Front. Cell Dev. Biol.* **10**, 895521 (2022).
47. S. Sano, S. Inoue, Y. Tanabe, C. Sumiya, S. Koike, *Science* **129**, 275–276 (1959).
48. T. Shetty, K. Sishtla, B. Park, M. J. Repass, T. W. Corson, *iScience* **23**, 101391 (2020).
49. M. Bulla, M. Šálek, A. G. Gosler, *Auk* **129**, 26–35 (2012).
50. C. Alonso-Alvarez *et al.*, *Ecol. Lett.* **7**, 363–368 (2004).
51. M.-Y. Lu, W.-W. Wang, G.-H. Qi, L. Xu, J. Wang, *Poult. Sci.* **100**, 100811 (2021).
52. S. Y. Kim, H. T. Kang, H. R. Choi, S. C. Park, *Exp. Mol. Med.* **43**, 15–23 (2011).
53. A. S. Baras, A. Solomon, R. Davidson, C. A. Moskaluk, *Lab. Invest.* **91**, 1170–1180 (2011).
54. S. S. Shinde, S. Sharma, L. Teekas, A. Sharma, N. Vijay, *Sci. Rep.* **11**, 24437 (2021).
55. J. Hu *et al.*, *Theriogenology* **232**, 87–95 (2025).
56. C. Yang *et al.*, *PLOS ONE* **5**, e10816 (2010).
57. J. N. Wolff, E. D. Ladoukakis, J. A. Enríquez, D. K. Dowling, *Philos. Trans. R. Soc. London Ser. B* **369**, 20130443 (2014).
58. L. Chen *et al.*, *PLOS Genet.* **16**, e1009119 (2020).
59. Z. Wang *et al.*, *PLOS Genet.* **9**, e1003183 (2013).
60. S. Ito, M. Tsudzuki, M. Komori, M. Mizutani, *J. Hered.* **84**, 145–147 (1993).
61. G. H. Kunz, *Naumannia Arch. Für Ornithol. Vorzugsweise Eur.* **1**, 51–53 (1850).
62. E. C. Collias, *Auk* **110**, 683–692 (1993).
63. J. Merondun, merondun/cuculus_host: Revisions: August 2025 v2, version v1.3. Zenodo (2025); https://doi.org/10.5281/zenodo.16816856.

ACKNOWLEDGMENTS

We thank all the field biologists who assisted with field work and blood collection, including G. Rudolfsen, I. Hafstad, A. Antonov (the late), J. R. Vikan, L. Wang, C. Yang, B. Zhou, G. Zhong, K. Schulze-Hagen, J. Berstad, and M. Požgayová, as well as N. Schäffer for additional support. **Funding:** Computing was performed on the BioHPC hosted at Leibniz Rechenzentrum Munich funded by German Research Foundation grant INST 86/2050-1 FUGG (J.B.W.W.). This work was funded by LMU Munich (J.B.W.W.); German Research Foundation grant WO 1426/2-1 (J.B.W.W.); The Research Council of Norway grant 151641/177709/218144 (F.F. and B.G.S.); Russian Federation grant 122021000096-7 (M.G.); National Natural Science Foundation of China grants 32270526 and 32470513 (W.L.); and National Key R&D Program of China grant 2023YFF1304600 (W.L.). **Ethical approval:** All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. Permission for blood sampling of wild common cuckoos was granted in the Czech Republic by the Hodonín Municipality Office (MUHOCJ 14306/2016/OŽP) with experimental design approved by the Czech Ministry of the Environment (3030/ENV/17–169/630/17). Collection of blood samples complied with the legal regulations of Bulgaria. Blood sampling and experimental procedures in China were in accordance with the Animal Research Ethics Committee of Hainan Provincial Education Centre for Ecology and Environment, Hainan Normal University (no. HNECEE-2012-002 and HNECEE-2012-003). In the UK, loose feathers from common cuckoos were collected as a by-product of tagging work carried out under license from the Special Marks Technical Panel of the British Trust for Ornithology Ringing Scheme, operating on behalf of the UK government's Home Office. Appropriate licenses were obtained for sampling in Finland (ESAVI/12343/2020). Permission for sampling of blood from wild common and oriental cuckoos in Russia in compliance with the Federal Law of Russian Federation no. 498-ФЗ, as ensured by the Research Bioethics Committee of the Institute of Plant and Animal Ecology, Russian Academy of Sciences. Necessary compliance with the Nagoya protocol was ensured for each country from which genetic resources were obtained. **Author contributions:** Contributions are listed in alphabetical order. Conceptualization: J.B.W.W., F.F.; Methodology: J.B.W.W., J.M., G.L.O.; Investigation: J.B.W.W., J.M.; Formal analysis: J.M.; Visualization: J.M.; Funding acquisition: J.B.W.W.; Project administration: J.B.W.W., G.L.O.; Resources: B.G.S., C.H., C.M., G.B., J.R., K.T., M.G., M.H., M.M., M.Š., P.A., P.P., S.M., V.B., V.F., W.L., Y.R.; Supervision: J.B.W.W.; Writing – original draft: J.M.; Writing – review & editing: J.B.W.W., J.M. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All associated sequence data are uploaded to the NCBI Sequence Read Archive (SRA) with biosample and run accessions indicated in table S1 under Bioproject PRJNA614488. All bioinformatic code in markdown .html and .md formats necessary to reproduce the findings presented here are available on GitHub (https://github.com/EvoBioWolf/CUCKOO_gentes) and are perpetually archived via Zenodo (*63*). **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. This research was funded in whole or in part by the German Research Foundation (INST 86/2050-1 FUGG and WO 1426/2-1); as required, the author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adt9355
Materials and Methods; Tables S1 to S7; Figs. S1 to S32; References (*64*–*126*);
MDAR Reproducibility Checklist

Submitted 18 October 2024; resubmitted 9 May 2025; accepted 2 September 2025

10.1126/science.adt9355

Who's the Top Employer for 2025?



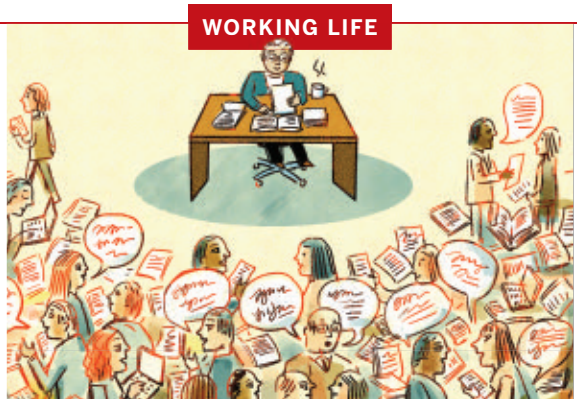
Science 2025 TOP EMPLOYERS

Science Careers' annual survey reveals the top companies in biotech & pharma voted on by *Science* readers.

Explore these highly-rated employers in our interactive experience:
sciencecareers.org/topemployers

ScienceCareers.org

ScienceCareers



Adjusting to retirement

Tod F. Stuessy

After more than 40 years in academia, I felt optimistic about the path ahead when I retired in 2012. I intended to continue working, focusing on research and writing now that I no longer had administrative and teaching responsibilities. But as I downsized my office and changed my email signature to Emeritus Professor, I was unprepared for the reality of retirement. I found it difficult to adjust to no longer holding decision-making roles. Colleagues were less likely to seek my opinions than before, which ate at my sense of self-worth and left me feeling that my ideas were no longer valued. Regaining my equilibrium and appreciating the positive aspects of retirement took months.

I had leadership roles throughout my career. After spending 27 years at a U.S. university, I pivoted to working as the deputy director of a natural history museum for 2 years followed by 15 years as a faculty member at a European university. In these positions, I had a range of decision-making responsibilities—departmental chair, botanical garden director, journal editor, secretary-general for an international scientific organization.

As I reached 68, I was ready to let those roles go. After so many cycles of dealing with administrative hurdles and other challenges, I had lost some of my drive. It seemed appropriate for more junior faculty to step into those shoes. But I hadn't fully processed what that might mean for me.

I was excited to have workdays free of the meetings and classes that broke up my days as a professor. I expected to have more time to finish manuscripts, start collaborative projects, and dive into book writing. That turned out to be true. But it was hard to ignore that meetings were being scheduled to which I was not invited. I was no longer an essential cog in the system.

As time wore on, I also had a nagging sense of being left behind scientifically, standing beside the track after the train had left the station. I continued to attend national and international conferences, but fewer than before. I didn't always hear about new developments, and I drifted away from old friends from other universities.

On my own campus, where I had a small office, the lack of daily chats in the laboratory with students and colleagues and discussions in the classroom left me feeling isolated. Retirement is a haven for the solitary worker, and working alone has never been a problem for me. But I did find that when bad news struck—such as learning that a paper had been rejected—I had no class later that day where a successful lecture could recharge my self-confidence.

Slowly, though, I began to ease into my new rhythm. Structure helped. I gave myself deadlines for completing manuscripts, and I set a schedule to get to the office by 10 a.m. each morning, which stimulated me to get out of bed and keep going. I also learned to see my isolation as a privilege rather than a penalty. In addition to being able to complete shorter manuscripts, I could focus on larger projects, such as review papers and books, which have proved very satisfying. Although retirement means I have lost influence with university decision-making, having the time to step back and take a broad, synthetic view of my field more than makes up for it.

For me, success in retirement has involved mapping out a realistic plan that is both enjoyable and academically fulfilling. When I do attend a conference, I go with reduced expectations—not to capture all the newest advances, but to learn about broad new trends and catch up with as many old friends as possible. The good news is that I am no longer being constantly evaluated by others—I am only being evaluated by myself. I have learned to keep pushing forward, at a more relaxed pace, and working with, rather than fighting against, these new challenges. □

Tod F. Stuessy is an emeritus professor at Ohio State University and the University of Vienna. Do you have an interesting career story to share? Send it to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER


IT'S NOT
JUST A JOB.
IT'S A CALLING.
Find your next job at ScienceCareers.org
For anyone who's looking to get ahead in — or just plain get into — science, there's no better, more trusted resource or authority on the subject than *Science* Careers. Here you'll find opportunities and savvy advice across all disciplines and levels. There's no shortage of global problems today that science can't solve. Be part of the solution.
ScienceCareers
FROM THE JOURNAL SCIENCE
AAAS



Pushing the Boundaries of Knowledge

As AAAS's first multidisciplinary, open access journal, *Science Advances* publishes research that reflects the selectivity of high impact, innovative research you expect from the *Science* family of journals, published in an open access format to serve a vast and growing global audience. Check out the latest findings or learn how to submit your research: **ScienceAdvances.org**

Science Advances
AAAS

GOLD OPEN ACCESS, DIGITAL, AND FREE TO ALL READERS