**Is abandoned farmland a boon for biodiversity?** p. 18

**Canadian forests are burning more** pp. 22 & 91

**Migrating bats surf warm fronts** pp. 25 & 97

Science
$15
3 JANUARY 2025
science.org
AAAS

# ALL ABOUT TIMING

Mating type evolution in walnuts and pecans p. 40

# CONTENTS

3 JANUARY 2025
VOLUME 387
ISSUE 6729

18 In Tyurkmen, Bulgaria, storks and other animals are reclaiming land abandoned by people.

# NEWS

## IN BRIEF

## IN DEPTH

## FEATURES

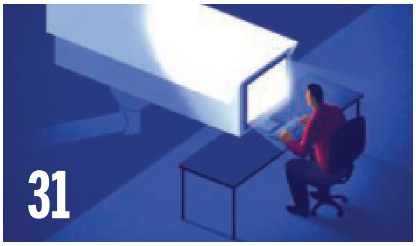
# INSIGHTS

## PERSPECTIVES

## POLICY FORUM

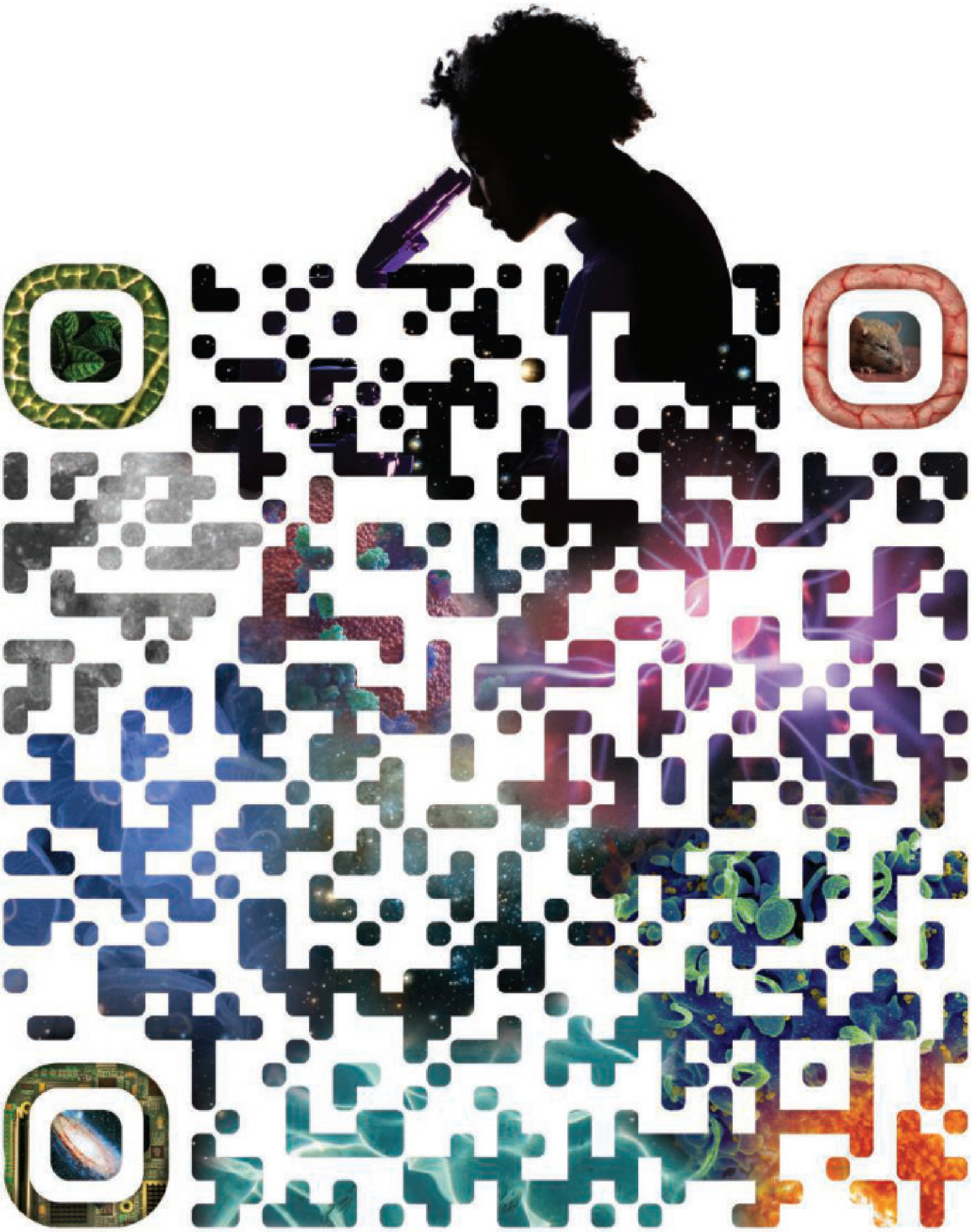
## BOOKS ET AL.

## LETTERS



31

CREDITS: (PHOTO) MALKOLM BOOTHROYD; (ILLUSTRATION) DAVIDE BONAZZI/ISPOT STOCK



# RESEARCH

## IN BRIEF

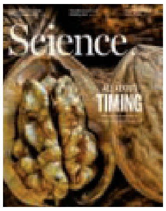
## REVIEW

## RESEARCH ARTICLES


25 & 97

## DEPARTMENTS

## ON THE COVER



Pollination in walnuts (such as this Persian walnut, *Juglans regia*) and pecans occurs through a genetically controlled temporal flowering dimorphism. Two distinct morphs alternate their phases of male and female flowering across the season, promoting outcrossing that has maintained a stable equilibrium throughout tens of millions of years of evolution. See page 40. *Photo: Philippe Clement/NPL/Minden Pictures*

PHOTO: PAUL HOBSON/NPL/MINDEN PICTURES

**SCIENCE (ISSN 0036-8075) is published weekly on Friday, 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. 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.

Science AAAS

**Editor-in-Chief** Holden Thorp, hthorp@aaas.org

**Executive Editor** Valda Vinson

**Editor, Research** Jake S. Yeston **Editor, Insights** Lisa D. Chong **Deputy Executive Editor** Lauren Kmec

**DEPUTY EDITORS** Stella M. Hurtley (UK), Phillip D. Szuromi, Sacha Vignieri **SR. EDITORS** Caroline Ash (UK), Michael A. Funk, Di Jiang, Priscilla N. Kelly, Marc S. Lavine (Canada), Sarah Lempriere (UK), Mattia Maroso, Yevgeniya Nusinovich, Ian S. Osborne (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 **SR. EDITORIAL MANAGERS** Carolyn Kyle, Beverly Shields **SR. PROGRAM ASSOCIATE** Maryrose Madrid **EDITORIAL ASSOCIATES** Aneera Dobbins, Joi S. Granger, Lisa Johnson, Anita Wynn **SR. EDITORIAL COORDINATORS** Jeffrey Hearn, Alexander Kief, Ronmel Navas, Jerry Richardson, Isabel Schnaidt, Alice Whaley (UK), Brian White **EDITORIAL COORDINATORS** Clair Goodhead (UK), Kat Kirkman, Samantha Price **ADMINISTRATIVE COORDINATOR** Karalee P. Rogers **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI OFFICE MANAGER** Victoria Smith **ASI SR. OFFICE ADMINISTRATORS** Dawn Titheridge (UK), Jessica Waldock (UK) **COMMUNICATIONS DIRECTOR** Meagan Phelan **DEPUTY DIRECTOR** Matthew Wright **SENIOR WRITERS** Walter Beckwith, Joseph Cariz **WRITERS** Abigail Eisenstadt, 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, Shraddha Chakradhar, Elizabeth Culotta, Martin Enserink, David Grimm, Eric Hand, Kelly Servick, Matt Warren (Europe) **SR. CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Michael Price **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Phie Jacobs, Jocelyn Kaiser, Rodrigo Pérez Ortega (Mexico City), Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **CONTRIBUTING CORRESPONDENTS** Vaishnavi Chandrashekhar, Dan Charles, Warren Cornwall, Andrew Curry (Berlin), Ann Gibbons, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Virginia Morell, Dennis Normile (Tokyo), Cathleen O'Grady, Elisabeth Pain (Careers), Charles Piller, Richard Stone (Senior Asia Correspondent), Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Katie Langin (Associate Editor) **INTERNS** Hannah Richter **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 GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **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 SYSTEMS AND FINANCIAL ANALYSIS** Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **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 **SENIOR CONTENT & PUBLISHING SYSTEMS SPECIALIST** Jacob Hedrick **SENIOR PRODUCTION SPECIALIST** Kristin Wowk **PRODUCTION SPECIALISTS** Kelsey Cartelli, Audrey Diggs **DIGITAL PRODUCTION MANAGER** Lisa Stanford **SENIOR DIGITAL ADVERTISING SPECIALIST** Kimberley Oster **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SR. TRAFFIC ASSOCIATE** Christine Hall **SPECIAL PROJECTS ASSOCIATE** Shantel Agnew

**MARKETING DIRECTOR** Sharice Collins **ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT** Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **ASSOCIATE DIRECTOR, MARKETING SYSTEMS & STRATEGY** Aimee Aponte **JOURNALS MARKETING MANAGER** Shawana Arnold **MARKETING ASSOCIATES** Ashley Hylton, Lorena Chirinos Rodriguez, Jenna Voris **SENIOR DESIGNER** Kim Huynh

**DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING** Erika Gebel Berg **ASSOCIATE EDITOR, CUSTOM PUBLISHING** Jackie Oberst **PROJECT MANAGER** Melissa Collins

**DIRECTOR, PRODUCT & PUBLISHING DEVELOPMENT** Chris Reid **DIRECTOR, PRODUCT MANAGEMENT** Kris Bishop **PRODUCT DEVELOPMENT MANAGER** Scott Chernoff **PUBLISHING PLATFORM MANAGER** Jessica Loayza **SR. PRODUCT ASSOCIATE** Robert Koepke **PRODUCT ASSOCIATES** Caroline Breul, Anne Mason

**ASSOCIATE DIRECTOR, INSTITUTIONAL LICENSING MARKETING** Kess Knight **BUSINESS DEVELOPMENT MANAGER** Rasmus Andersen **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, US ADVERTISING** Stephanie O'Connor **US MID WEST, MID ATLANTIC AND SOUTH EAST SALES** Chris Hoag **US WEST COAST SALES** Lynne Stickrod **ASSOCIATE DIRECTOR, INTERNATIONAL** Roger Goncalves **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/products-services
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** Gilda A. Barabino
**PRESIDENT** Keith R. Yamamoto
**PRESIDENT-ELECT** Willie E. May
**TREASURER** Carolyn N. Ainslie
**CHIEF EXECUTIVE OFFICER**
Sudip Parikh
**BOARD** Cynthia M. Beall
Janine Austin Clayton
Kaye Husbands Fealing
Kathleen Hall Jamieson
Jane Maienschein
Robert B. Millard
Babak Parviz
William D. Provine
Juan S. Ramírez Lugo
Susan M. Rosenberg
Vassiliki Betty Smocovitis

EDITORIAL

# Steady going in 2025

As the calendar turns to 2025, the unflagging pace of advancement in science and technology promises to break new ground in areas as profound as humanity's relationship with artificial intelligence (AI), gene editing in medicine and agriculture, climate change mitigation, and more. Meanwhile, as research across all fields becomes more complex, the challenge of maintaining research integrity and the reliability of scientific findings becomes more important than ever for informing public policies and bolstering trust in science. All organizations in the scientific enterprise have a responsibility to ensure that sound practices and ethical standards are at the heart of their operations and missions. This includes scientific journals.

Here at *Science*, we are making changes focused on strengthening the scientific record, helping authors submit papers with complete and robust data, and recognizing experts for their role in the peer review and publication process. Against new political backdrops around the world in 2025, there may be changes that affect the endeavors of science. Regardless, by continuing to strengthen our policies and base our practices on solid scientific principles, *Science* will remain a critical tool for communication and progress for the scientific community.

As in the past, the *Science* family of journals cannot achieve its goal of publishing excellent science without rigorous peer review. To acknowledge the invaluable contribution of anonymous peer reviewers who volunteer in this process, we are collaborating with the nonprofit organization ORCID (Open Researcher and Contributor ID). This partnership allows reviewers to establish a record of their work for *Science* journals without identifying the paper that was reviewed. We also encourage reviewers to involve trainees from their laboratory as co-reviewers and now have a process that indicates such participation in *Science*'s database.

Following our announcement last year, the *Science* family of journals began using the Proofig software for image analysis across papers that are peer-reviewed. This tool enables the identification of image duplications and manipulations, image plagiarism, and AI-generated imagery. Since January 2024, we have analyzed more than 2000 papers and documented 23 with issues that prompted a request for explanations and raw data from the authors, and occasionally a request to reviewers to consider this new information. Most issues turned out to be errors in figure assembly and could be remedied before publication. However, two papers were rejected because the authors could not provide a reasonable explanation for the irregularities, undermining confidence in the data. This experience generally points to the need for careful data management in laboratories to prevent errors during figure compilation.

Building on a *Science* editorial from November 2023, the journal has strengthened its corrections policy for greater transparency. As a result, *Science* published more erratum notices in 2024 than in previous years (30 versus 12, 4, and 9 in 2021, 2022, and 2023, respectively). These now appear prominently at the top of the related paper. When an error is uncovered, it is important to issue a correction as swiftly as possible—not to assign blame, but to maintain an accurate scientific record.

> "...the integrity of science across the enterprise should never be compromised."

Although we continue to encourage readers with concerns about the integrity of a paper published in a *Science* journal to contact us at science_data@aaas.org, we have also taken a more concentrated approach by routinely monitoring the community website PubPeer and following up on concerns raised there as well as in other places. At PubPeer, users can discuss scientific research after it has been published. Sometimes authors and institutions differ on whose responsibility it is to address comments at PubPeer. We are committed to examining any concerns that may affect the substance of our papers. In the past 18 months, the *Science* journals have been writing to authors about PubPeer comments to determine whether further follow-up is required.

We also want to inform readers about another change. *Science* has had a long tradition of publishing Retrospective articles on prominent scientists who have passed away. Because of resource constraints, we have made the difficult decision to discontinue these articles so that we can focus on *Science*'s core scientific content. We apologize to readers who have enjoyed this feature and hope they will feel compensated by our continuing efforts to publish scientific journals of the highest quality.

Scientific journals must ensure the transparency, accountability, and credibility of science. Changes in the culture and expectations of both publishers and researchers will always arise, but the integrity of science across the enterprise should never be compromised.

**–H. Holden Thorp, Valda Vinson, Lauren Kmec**

**H. Holden Thorp**
Editor-in-Chief of the *Science* journals.
hthorp@aaas.org

**Valda Vinson**
Executive Editor of the *Science* journals.
vvinson@aaas.org

**Lauren Kmec**
Deputy Executive Editor of the *Science* journals. lkmec@aaas.org

10.1126/science.adv4633

# NEWS

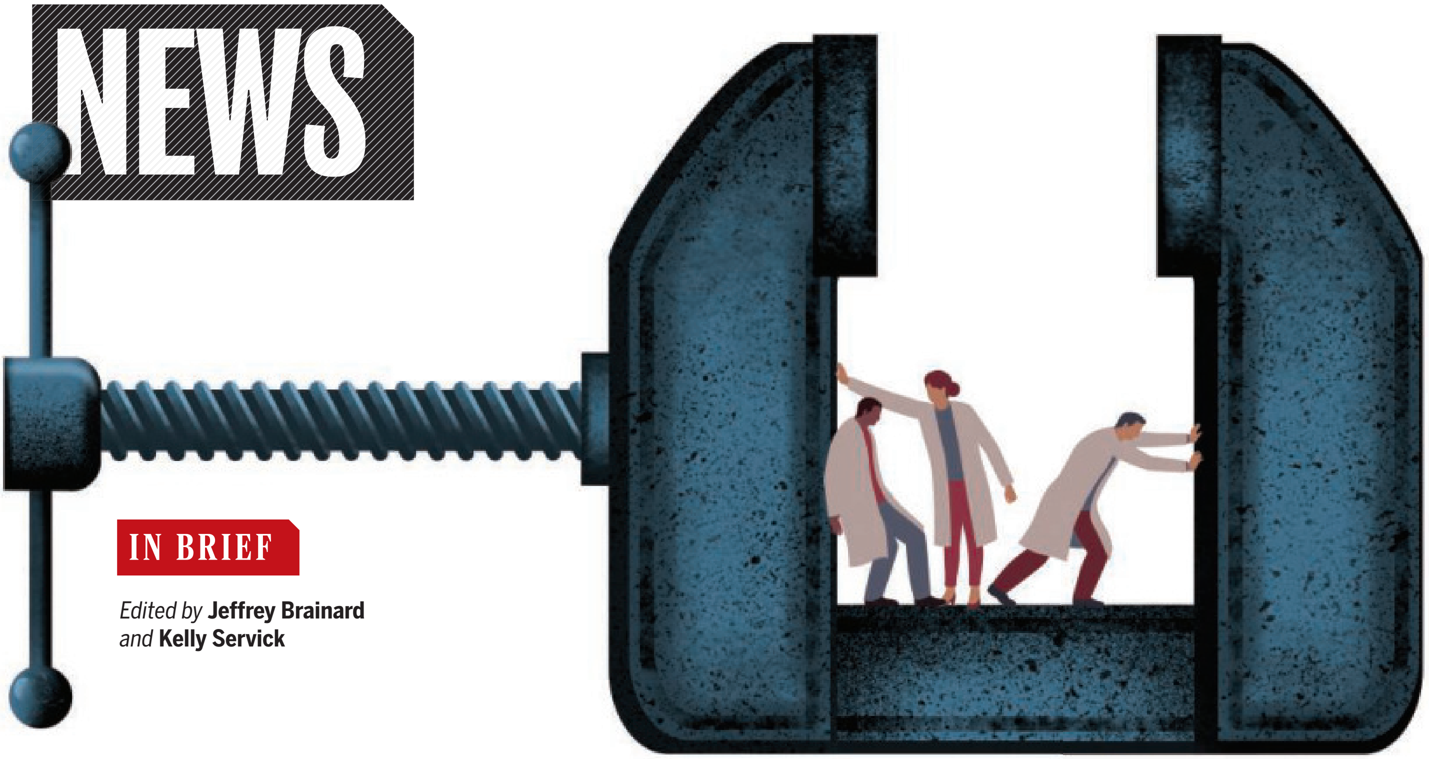


ILLUSTRATION: DAVIDE BONAZZI/SALZMAN ART

## IN BRIEF

*Edited by* **Jeffrey Brainard**
*and* **Kelly Servick**

AREAS TO WATCH: OUR FORECAST FOR 2025

# Scientists brace for Trump's second act

Will his bark be worse than his bite? President-elect Donald Trump doesn't take office until 20 January, but his repeated promises to shrink the government and slash spending have already put the U.S. research community on edge. And several of his nominees have pledged to shake up the science agencies they are expected to oversee, including the National Institutes of Health, the biomedical colossus under fire for its role in combating the COVID-19 pandemic. There's also angst about possible political interference in grantmaking and erosion of scientific integrity policies across government. Research on climate change, conservation, and renewable energy sources, which fell out of favor in Trump's first administration, is again expected to be squeezed, along with new targets such as programs aimed at creating a more inclusive scientific workforce. Republican members of Congress, who will control both of its chambers, are expected to continue a wide-ranging investigation of U.S. higher education. But House of Representatives Speaker Mike Johnson (R–LA) could have difficulty managing his party's razor-thin majority. Legislators are also protective of their constitutional authority over federal spending, and many programs that Trump has criticized enjoy bipartisan support. In the rest of this section, *Science*'s News staff previews other research and policy areas likely to make headlines this year.

## H5N1 influenza's evolving threat

**INFECTIOUS DISEASES** | Two major questions loom for the influenza field this year. One is whether the United States can get rid of the now-massive outbreak in dairy cattle of an H5N1 avian flu strain that likely started in late 2023. This year, bulk milk testing, mandated under a federal order last month, may help identify infected herds earlier and curb the spread of the virus. Results from vaccine trials in cows are also expected, and scientists hope to better understand why most people sickened by the H5N1 variant affecting cattle suffered only a mild eye infection, whereas those infected after contact with birds typically have more serious disease. The other big question: Will its spread in mammals—and frequent infections in humans from both birds and cows—give H5N1 more avenues to evolve, eventually sparking a pandemic? Scientists are on the alert for signs that the virus is adapting to humans and becoming transmissible between people.

## Peak emissions at last?

**CLIMATE CHANGE** | Scientists are hopeful this year could mark a critical turning point in the fight against climate change: the year when global greenhouse gas emissions peak. Annual emissions, driven mostly by the combustion of fossil fuels, seem to have nearly plateaued, increasing by about 1 percentage point each year for the past 2 years and totaling 41.6 billion tons in 2024. The rapid rise of electric vehicles, renewable energy, and reforestation has jostled with the countervailing forces of the energy-gobbling data centers powering artificial intelligence and rebounding fuel demand after a pandemic lull. As the world's largest emitter, China, continues its aggressive push into renewables, many researchers suspect this year could finally see a long-awaited drop. But even if the world reaches that milestone, it is expected to take several decades before the world reaches "net zero" and emissions return to preindustrial levels. And even then, the warmth from existing carbon dioxide will linger for centuries.

## Lifestyle clues in ancient bones

**ARCHAEOLOGY |** In the coming year, scientists expect to make leaps in identifying chemical signatures in bones of long-dead people, offering new kinds of clues to their behavior. Physicians, forensic specialists, and biomedical researchers have long traced metabolites from the foods, beverages, and medicines people consume in samples of their hair, blood, urine, and saliva. Now, researchers are examining such metabolomic changes in well-preserved tissue such as bone. In 2024, archaeologists identified the chemical signatures left by tobacco smoking in the bones of hundreds of English people who lived between 1700 and 1855, offering a window into their health and social customs. Scientists hope to find more chemical hallmarks in other ancient bone samples, including those of our hominin ancestors. One study now underway is looking at metabolites in Egyptian mummies to better understand how diseases such as tuberculosis and plague impacted various levels of society. Other research aims to dig deeper into how ancient people made themselves merry, tracing metabolites of mood-altering substances such as alcohol, tobacco, and coca.

## Shots may drive down malaria

**GLOBAL HEALTH |** Public health experts and policymakers hope to see a noticeable dent this year in malaria cases and deaths among children in the 17 countries where two malaria vaccines have been rolled out on a large scale. More than a dozen countries in sub-Saharan Africa added the shots to their routine childhood vaccination schedules this year; GAVI, the Vaccine Alliance estimates that in 2024 about 5 million children received at least one of the recommended four doses, given over 12 months. This year, GAVI aims to reach 14 million children in 25 countries. Although comprehensive data on malaria cases, hospitalizations, and deaths are notoriously tricky to collect, researchers say it should be possible to see disease rates go down in regions where children have been vaccinated. Pilot programs in Ghana, Kenya, and Malawi that reached more than 2 million children between 2019 and 2023 found that vaccination reduced the number of children admitted to hospitals with severe malaria by nearly one-third and reduced overall mortality by 13%.

## Diversity initiatives under fire

**DIVERSITY |** U.S. laws aimed at restricting diversity, equity, and inclusion (DEI) initiatives in higher education and research institutions are expected to ramp up this year, building on the 14 laws already enacted in 12 states since 2023—and a Republican-controlled Congress may also take up the issue. Some of the laws passed to date restrict use of federal or state funds for DEI activities; many eliminate DEI offices and positions at public universities and prohibit employers from requiring diversity training for staff. In response, some universities have adjusted their policies and halted some DEI efforts. For instance, in December 2024, the University of Michigan ceased requiring diversity statements in faculty hiring and promotion packages. President-elect Donald Trump has also vowed to eliminate DEI administrators at educational institutions and cut their funding if they teach critical race theory and gender education. Together, such moves could influence the distribution of DEI research funding, recruitment of diverse populations to clinical trials, and universities' efforts to attract and retain scientists from underrepresented groups.

## Charting the changing sky

**ASTRONOMY |** Making its debut this year will be the Vera C. Rubin Observatory, a U.S.-funded survey telescope poised to transform many areas of astronomy. Instead of zooming in on distant objects, every 3 days the observatory's 8.4-meter main mirror will record changes in the entire sky visible from its mountaintop perch in Chile. To achieve this, the Rubin Observatory has the largest camera ever built, along with a structure compact and agile enough to shift positions in just 5 seconds. By comparing images with earlier ones, it can pinpoint celestial objects that have moved, changed brightness, or suddenly appeared or disappeared, such as previously unknown comets and asteroids, exploding stars in distant galaxies, or, perhaps, Planet Nine, a hypothesized giant planet beyond Neptune. The Rubin Observatory will find up to 10 million such "transients" every night. Other tasks include tracking the history of the dark energy that is accelerating the universe's expansion, and charting how dark matter has shaped the evolution of galaxies.

## Will new guidelines trim the beef?

**NUTRITION |** A battle is simmering over whether the U.S. government will propose limiting consumption of red meat as part of a recommended healthy diet. About every 5 years, the Departments of Agriculture and Health and Human Services jointly issue dietary guidelines that influence food labels and the meals prepared by schools and other institutions. They are expected to issue new guidelines this year. Last month, an advisory committee of nutritionists and public health experts finished its recommendations for the guidelines, calling for reducing consumption of red and processed meats in favor of beans and other plant protein. This is mainstream nutritional advice but it drew swift criticism from industry and members of Congress from farm states. President-elect Donald Trump's incoming administration may reject the new advice; his first administration's 2020 guidelines rejected a previous advisory group's suggested limits on sugar and alcohol consumption.

The Vera C. Rubin Observatory's massive camera will produce a 10-year time-lapse movie of the universe.

PHOTO: ALIRO PIZARRO DÍAZ/RUBIN OBSERVATORY/NSF/AURA

PHOTO: HECTOR RETAMAL/AFP VIA GETTY IMAGES

INFECTIOUS DISEASES

# Learning from a pandemic many are forgetting

Five years after COVID-19 surfaced, scientists reflect and look ahead to the next threat

*By* **Jon Cohen**, *in Awija, Japan*

The COVID-19 pandemic, as best as we can tell, took more than 20 million lives, cost $16 trillion, kept 1.6 billion children out of school, and pushed some 130 million people into poverty. And it's not over: In October 2024, at least 1000 people died from COVID-19 each week, and that's relying only on data, some questionable, from the 34 countries that still report deaths to the World Health Organization (WHO). Last month, at a 4-day meeting here on preventing future pandemics, WHO epidemiologist Maria Van Kerkhove ticked off those figures with exasperation. "The world I live in right now, no one wants to talk about COVID-19," she told the gathering. "Everyone is acting as though this pandemic didn't really happen."

Yet 5 years after a coronavirus dubbed SARS-CoV-2 first surfaced in Wuhan, China, scientists are still intensively trying to make sense of COVID-19. "We would each have to read over 240 papers every single day to actually keep up with all of the [COVID-19] literature that's come out" in 2024, Cherilyn Sirois, an editor at *Cell*, noted.

Despite the flood of insights into the behavior of the virus and how to prevent it from causing harm, many at the meeting worried the world has turned a blind eye to the lessons learned from the pandemic. "I feel this massive gravitational pull to go back to what we were doing before," Van Kerkhove said. "There's no way we should be going back."

Even more concerning to some at the conference, many countries have become hostile toward research into other potential pandemic agents, much of the anger stemming from an unproven assertion that SARS-CoV-2 leaked from a lab. "There's been massive public and political backlash against the virology community and public health in general, so we may be worse off now locally than we were prior to the pandemic," said virologist Ralph Baric of the University of North Carolina at Chapel Hill. Baric was recently accused by Robert Redfield, former head of the U.S. Centers for Disease Control and Prevention, of being the "scientific mastermind" of a supposed effort to engineer the virus.

The conference, held under the auspices of Cold Spring Harbor Asia, brought together 140 researchers and health officials from 17 countries to discuss everything from the origin of the pandemic to SARS-CoV-2's mutational patterns, new treatments, and creative vaccine strategies to fend off future threats. "One of the big reasons we wanted to hold this conference is because we couldn't meet in person during the pandemic," said virologist Kei Sato of the University of Tokyo, an organizer. He also hoped the location would attract infectious disease scientists from China, who have had limited interactions with the global research community since 2020.

Some 20 scientists from China attended, but two of the country's most prominent COVID-19 researchers were notable no-shows. Shi Zhengli, who studied bat coronaviruses at the Wuhan Institute of Virology (WIV) and became the focus of intense criticism by people who suspect SARS-CoV-2 leaked from her lab, only gave a prerecorded video talk—on the sequencing of other coronavirus genomes—despite being a co-organizer of the event. So did Chinese Academy of Sciences virologist George Gao, former head of China's Center for Disease Control and Prevention. Sato suspects the Chinese government would not let either attend. Shi, now with the Guangzhou Laboratory, and Gao both declined to explain their absence to *Science*.

A Chinese scientist did offer one of the most provocative presentations. Immunologist Yunlong Cao of Peking University, another organizer of the meeting, noted that

After months of no contact because of COVID-19, a woman in 2020 found a way to hug her grandmother.

the "extraordinary viral evolutionary speed" of SARS-CoV-2 not only means fresh variants are "continuously causing reinfections," but that antibody treatments and vaccines quickly lose effectiveness. None of the first approved monoclonal antibodies and vaccines work against current circulating strains.

Cao noted that only two of 140 antibodies his lab identified in early 2020 as able to neutralize the first variant of SARS-CoV-2 could protect against the virus in circulation 2 years later. "The only solution to this problem," he said, "is if we can do accurate predictions about viral evolution" to assess which antibodies will retain their powers.

Cao's group recently identified an antibody, dubbed SA55, that he predicts will stymie whatever SARS-CoV-2 variants evolve for at least two more years. His team began by drawing blood from an unusual cohort of 28 people who nearly 2 decades ago contracted and recovered from severe acute respiratory syndrome (SARS), an illness caused by another coronavirus, and then during the pandemic received vaccines against SARS-CoV-2. The researchers isolated some 13,000 memory B cells and screened the antibodies they made for the ability to neutralize both coronaviruses and several relatives found in bats and other species. SA55 stood out as a superstar.

Sinovac Biotech, one of China's largest pharmaceutical companies, has tested a nasal spray containing SA55 as a preventive. In clinical trials, researchers gave it at least two times each day to participants who came in contact with infected people at home or work. The spray had about 80% efficacy at preventing infections, according to still-unpublished work, Cao says. Under "compassionate use" regulations, some 300,000 people in China have already received this spray and Sinovac plans a large phase 3 efficacy study.

Accurately predicting how viruses will evolve could also allow vaccinemakers to design more durable products. Computational biologist David Robertson of the University of Glasgow is part of a growing effort to peer into the future of SARS-CoV-2 with artificial intelligence (AI). His team is using protein language models—which convert genetic sequences of the virus into predicted protein structures—to map "evolutionary landscapes" that indicate how viral proteins might mutate and still retain their ability to infect new hosts and copy themselves. Ultimately, Robertson says, the models present the "exciting" possibility that they could guide design of vaccines that produce antibodies able to thwart a wide range of potential variants.

Many speakers emphasized that Omicron and other SARS-CoV-2 "variants of concern" have evolved in people who have weakened immune systems and cannot quickly clear infections. "A compromised host comes along and a weirdo virus comes out," said University of Sydney evolutionary biologist Edward Holmes. He calls AI prediction of virus evolution an "amazing tool," but cautions it has "a good way to go" in forecasting the variants that will emerge during ever changing pandemic environments.

Presenters also explored how to protect against other threatening coronaviruses. Molecular pharmacologist Gurpreet Brar of the Coalition for Epidemic Preparedness Innovations (CEPI) described how the nonprofit is funding development of vaccines for nine coronaviruses that have been found in mink, pigs, cattle, dogs, camels, and bats. "These are ones that have a high risk of spillover, and if they were to jump into humans we'd have a big problem," Brar said.

Major mysteries remain about what SARS-CoV-2 is doing today, and where it came from. There's no consensus on how the virus produces Long Covid, the debilitating symptoms that have afflicted millions after their infections have seemingly been cleared, or how to treat or prevent the condition. And efforts to unravel the pandemic's origin have largely stalled.

> **"Everyone is acting as though this pandemic didn't really happen."**
>
> **Maria Van Kerkhove**, World Health Organization

Meeting organizers, concerned that people angered by the idea SARS-CoV-2 leaked from a lab would crash the gathering and might harm invited scientists, hired extra security guards. But only Jonathan Latham, a virologist with the Bioscience Resource Project, made the case publicly for a lab leak. He presented a poster contending that SARS-CoV-2 came from tissue samples of copper miners, which WIV studied to determine what mysteriously sickened them in 2012.

Virologist Angela Rasmussen from the University of Saskatchewan challenged him in a heated confrontation, arguing that no evidence supports his theory. She later gave a talk that described her own attempt to find new information in a much-studied set of "environmental samples" collected between January and March 2020 from the Huanan Seafood Wholesale Market in Wuhan, which has been linked to many of the first COVID-19 cases in December 2019. "I'm slicing the salami ever thinner," Rasmussen said.

She was among a large contingent of researchers at the meeting who contend those samples and other evidence supports the theory that animals at the market carried SARS-CoV-2—and sparked the pandemic. Rasmussen reported that the market samples held animal genes that had been turned on by interferon, a response to viral infections. She ultimately concluded that raccoon dogs and greater hog badgers were the two most likely wild animals at the market to have been infected with SARS-CoV-2. But she acknowledged the limits of this analysis: "Spoiler alert: I have not found an infected animal."

Virologist Jesse Bloom of the Fred Hutchinson Cancer Center, who is not convinced the pandemic began at the market and has urged colleagues to remain open to the possibility of a lab leak, wasn't swayed by Rasmussen's new work. "There's still little actual information about the first human cases," Bloom says. "There's just not a lot of knowledge about what was really going on in Wuhan in late 2019."

Christian Drosten, a coronavirus researcher from the Charité Institute of Virology in Berlin who thinks the evidence strongly supports the market theory, decried the politicization of the origin debate. He was particularly exercised by a recent report from a Republican-led House of Representatives panel arguing that U.S. funding may have helped create SARS-CoV-2 at WIV before it leaked. "It's very clear that they are not only ignoring existing evidence, but they are falsifying the evidence that's on the table," Drosten said. "It's really surprising and puzzling that people at this meeting aren't speaking up. We will be quiet until we don't have a chance to speak anymore."

WHO's Scientific Advisory Group for the Origins of Novel Pathogens is expected to issue its own report on how the pandemic began in the next few weeks. But no one at the meeting anticipated major revelations. "I fully believe there's much more data that's out there that we don't have access to," says Van Kerkhove, who oversees the group. She knows of a Chinese database that has some 500 viral sequences from January and February 2020 that WHO cannot access. "The biggest question I have are the farms," she says, referring to the possibility that SARS-CoV-2 came from animals being bred to sell at markets for their meat.

As for the future, Van Kerkhove warns that the world is dropping its guard against novel pathogens. Infectious disease is "not a safe space to really be working in," she told *Science*. "Labs have been threatened. People have been threatened. Governments don't necessarily want to be the ones to say, 'Hey, we found something new.'" ■

Demonstrators at Damascus University drag a statue of Syria's late president, Hafez al-Assad, days after rebels seized the city last month.

PHOTO: OMAR HAJ KADOUR/AFP VIA GETTY IMAGES

CONFLICT

# Can war-torn Syria rebuild its scientific community?

## Syrian researchers around the world begin to plan for the nation's future, including fixing "miserable" universities

*By* **Richard Stone**

In 2012, when Syrian government forces started bombing the eastern city of Deir ez-Zor early in the country's civil war, conservation biologist Ahmad Aidek barely managed to escape his besieged neighborhood. In his apartment, he left behind dozens of rare specimens including the shell of an Eiselt's pond turtle and a mummified sand gazelle. Six years of bombardment "completely destroyed" most of the city, not to mention his precious specimens, Aidek says. "I lost them all."

But Aidek never lost his passion for science. Whereas many of Syria's best researchers sought refuge abroad during the 13-year civil war, Aidek stuck it out in Deir ez-Zor at one of the nation's few remaining science institutes, the General Commission for Scientific Agricultural Research. As the war ground on, he published articles with European co-authors on Syria's biological riches, including a survey of mammals to appear in *Zootaxa* this month. It became a lonely pursuit as Aidek watched the nation's academic standards crumble. "The state of Syrian universities is miserable," he says.

Now, the question is whether Syria's new leaders, who last month toppled President Bashar al-Assad's regime, can repair the damage. So far, rejuvenating Syria's academic research community is "not on the radar" of the Sunni paramilitary group Hayat Tahrir al-Sham that is consolidating power in Damascus, says Ahmad Sadiddin, a Syrian agricultural economist at the Food and Agriculture Organization of the United Nations who left his homeland in 2012. More urgent problems include easing the plight of millions of Syrians displaced by the war, many living in tents without access to proper nutrition, let alone classes. "Housing and food have to come first," says neuroscientist Amal Alachkar of the University of California, Irvine, who also left Syria in 2012.

But she is losing hope that some of her former students at the University of Aleppo, whom the government detained years ago, will emerge from regime prisons emptied after Assad's downfall. "Most likely they were tortured to death," she says.

Once the immediate crisis passes, Syria's long-term recovery "will rely on rebuilding higher education," Alachkar says. And with the backing of institutions that threw them lifelines years ago, she and other members of Syria's scientific diaspora are hatching plans to help beleaguered colleagues who stayed.

"The good news is that we've saved a lot of Syria's professors," says Allan Goodman, president of the Institute of International Education (IIE), which over the past decade handed out 203 fellowships to 111 Syrian scholars, securing havens for them at institutions in Europe, North America, and the Middle East. "They've practiced their craft while they've been in exile," Goodman says. "There is a real chance they can build [Syrian higher education] back better."

Syria may not have been an academic powerhouse under the former regime, but it had a few science oases. The most vaunted had malevolent aims: the Syrian Scientific Studies and Research Center (SSRC), a sprawling military R&D enterprise thought to employ as many as 20,000 people. Israel bombed SSRC's institutes, including chemical weapons labs, hours after Assad fled the country on 7 December 2024.

Civilian science had begun to hollow out years earlier. In 2012, the war forced the nonprofit International Center for Agricultural Research in the Dry Areas to relocate from Aleppo to Lebanon. Syria's universities, meanwhile, withered as sanctions imposed on the Assad regime strangled the economy. Professors at public universities earn about $80 a month—"absolutely not enough" to subsist on, Aidek says. Such parlous conditions have turned Syrian campuses into hotbeds of corruption and scientific fraud. "Many university professors take bribes from students to pass their courses," or earn advanced degrees, Aidek asserts. Theses, he says, are often riddled with "false and completely unrealistic information."

Under Assad, the handful of reputable researchers who stayed in Syria had to maintain low profiles. "For the safety of my colleagues there, I kept only minimal contact," Alachkar says. "Speaking with them could put their lives in risk."

Like many Syrian scientists who went abroad early in the civil war, Alachkar initially thought she would return to her homeland within months. Now, she intends to help Syria-based scientists and students from her perch in California. For starters, Alachkar says, she and others could update curricula and teach online courses. Down the road, she foresees taking trips to Syria to mentor students and help establish research centers.

IIE, for one, is eager to facilitate such efforts if the new regime allows and is planning a meeting of U.S. stakeholders this month to discuss how to support Syrian academia. "If the chips fall well, we can immediately get to work on rebuilding their higher ed," Goodman says.

Much will have to be refashioned from scratch, Aidek says. Although Syria's academic ranks include genuine scientific talent, he predicts the country will need "a whole new generation of honorable researchers for our universities to resemble those in the rest of the world." ■

PUBLISHING

# U.S. agencies roll out free paper-access policies

## NIH and DOE are first to act, with implementation by all set to begin by end of this year

*By* **Jeffrey Brainard**

Two years after President Joe Biden's administration shook up scientific publishing by calling for immediate free access to journal articles produced from federally funded research by the end of this year, the U.S. National Institutes of Health (NIH) and Department of Energy (DOE) have released their final plans for complying. The policies require grantees to post accepted, peer-reviewed manuscripts in each agency's public repository upon publication and immediately share project data, among other stipulations; more U.S. agencies are expected to follow soon. Satisfying the 2022 request from the White House's Office of Science and Technology Policy (OSTP), the policies could mark a major step forward for open science: Nine percent of the world's research papers have U.S.-funded authors. But universities are worried about logistics and costs of this openness, and many publishers are dismayed.

One way to satisfy the new requirement is gold open access (OA), in which authors or institutions pay an article-processing charge (APC) for making a paper immediately free to read. The fees average about $2000, and critics have called the model unsustainable. (The *Science* family of journals includes one that charges APCs; *Science*'s News department is editorially independent of the journals.)

An alternative that avoids APCs is immediate public access—green OA—in which authors or institutions upload an accepted manuscript, lacking polishing of the publisher's paywalled version, into a repository. Since 2013, some agencies have required a form of this but allowed publishers to embargo the public release for up to 12 months.

Last year, universities, academic organizations, faculty, and librarians called on agencies to adopt the zero-embargo green approach. But several publishers argued immediate green OA would remove the incentive for readers and institutions to pay subscription fees, which support journals' editorial processes. The policies from OSTP, NIH, and DOE do not endorse any one business model for zero-embargo public access and do not bar authors from paying an APC.

Under the new policies, small, nonprofit scientific societies could be particularly vulnerable. The American Society of Clinical Oncology (ASCO), for example, could take a hit: Up to 40% of the papers in its four subscription journals are authored by NIH grantees, says Angela Cochran, the society's vice president for publishing. "Librarians have said to me that they don't have any plans to drop subscriptions to journals just because a large portion of the content is available as green open access. But … there are financial realities that may change that answer over time."

NIH will need to mount an educational campaign pointing authors to the no-APC alternative, university officials told the agency in written comments about the draft policy. "Our researchers and librarians are already hearing incorrect information from some publishers who claim public-access policies require federally funded researchers to publish open access and pay an APC," wrote Karen Caputo of Case Western Reserve University.

> **"Zero-embargo green [open access] is coming on a large scale."**
>
> **Peter Suber**, Harvard University

Other university officials cited practical obstacles. Depositing content in NIH's repository, PubMed Central—including formatting an article and providing required metadata—is time-consuming. "This isn't like a person uploads a Word file or a PDF and they're done," Cochran adds. Currently, publishers often do this work. Universities also objected that if a grantee publishes a paper through paid OA after their NIH grant has expired, the agency will not foot the APC bill.

Some universities already require or encourage faculty members to deposit the accepted manuscript in a public, institutional repository, says Peter Suber of Harvard University, where he is a senior adviser on its public-access policies. "But a lot of people need hand holding" to comply, Suber says, and at some universities, "there aren't enough hands to do all the holding."

The policies also raise questions about copyright and who controls whether an embargo applies. NIH and DOE assert that work they fund is covered by a "government use license," authorized by an existing U.S. regulation, that supports zero-embargo depositing of grantees' papers—overriding contracts authors sign with publishers. NIH also asserts that the government license allows other uses of the texts, a stance that may limit publishers' ability to charge a fee for working with their articles. A use highly sought by outsiders is employing artificial intelligence (AI) to analyze the papers for research. Some publishers have already sold access to AI developers. Taylor and Francis, for example, signed a $10 million deal with Microsoft, and ASCO's revenue from such licensing matches what it earns from journal subscriptions, Cochran says.

Open-science advocates have long argued that such access should be free. The 2022 OSTP policy recommended that deposited manuscripts be "machine readable," and NIH's draft wording required that. Publishers objected, and the final NIH version omits "machine readable" but promises "usability" that is "consistent with copyright law."

Whether the new policies will survive scrutiny by Congress remains a wild card. An appropriations bill for this fiscal year in the House of Representatives bars funding for implementing the 2022 policy. The corresponding bill in the Senate directs OSTP "to instruct [agencies] not to limit grant recipients' ability to copyright, freely license, or control their works." Final resolution on the bills—which govern OSTP and the National Science Foundation but not NIH, DOE, or several other grantmaking agencies—is unlikely to come until after agencies have announced their final public-access policies.

President-elect Donald Trump could also kill the policies, as he has pledged to slash government regulations. But in 2020 his administration was close to issuing an executive order requiring agencies to implement a public-access policy similar to the new ones. Some publishers objected to it, and he left office without issuing the order.

Despite the political uncertainty, some are confident the die is cast. As Suber puts it, "Zero-embargo green is coming on a large scale." But grantees will still need time to adapt, says Brian Hitson, a DOE official who helped lead the development of that agency's policy. DOE plans to give grantees up to 2 years to fully comply. "This is not something that's going to happen overnight." ■

Chance Ward says animal remains such as this ancient bison skull should get thoughtful care in museum collections.

PHOTO: AMANDA HEIDT

ARCHAEOLOGY

# Care urged for animal remains tied to Indigenous peoples

## Recommendations stress tracing and respecting links between communities and animal remains such as bones

*By* **Amanda Heidt**, *in Boulder, Colorado*

As a master's student in museum studies at the University of Colorado (CU) Boulder, Chance Ward opened a drawer one day in 2021 and found a jumbled bag of horse bones, many broken into bits or ground into dust. Ward is Lakota and an enrolled citizen of the Cheyenne River Sioux Tribe, and he and other Lakota consider the horse sacred—"an animal I see as my brother," he says. "It was upsetting to see [so little] care for them." The experience was so jarring that for his thesis, Ward explored ways for museums to more thoughtfully care for animal remains that have cultural ties to Indigenous communities.

Now, his work forms the basis of a paper, due out this month in *Advances in Archaeological Practice*, in which scientists, tribal elders, and preservation officers outline recommendations for the care of animal remains in anthropological collections. Curators should closely study animal remains for any links with Indigenous nations or communities, and researchers should involve those groups when remains are culturally significant, Ward says. Although animals don't typically fall under the Native American Graves Protection and Repatriation Act (NAGPRA), which enables Indigenous peoples in the United States to reclaim human remains and funerary objects, he and others say curators should consider returning some remains.

The work feeds into a museum movement toward culturally informed curation and collaborative care of collections. "This is such a timely paper, given how much energy we've seen in the space of ethical curation in the last decade or so," says Alex Fitzpatrick, a zooarchaeologist at London's Science Museum. "I've always found a frustrating discrepancy in how people and animals are treated, so it's really exciting to see a group of authors get together and actually put up a plan."

Museums contain thousands of masks, jewelry, clothing, and other ceremonial items crafted from animal materials. And millions more animal specimens have never been inventoried, which "drives the true tally much, much higher," says co-author William Taylor, a zooarchaeologist at CU Boulder who was Ward's master's thesis adviser.

The glut of unprocessed material—dubbed a "curation crisis" by museum professionals—underpins the paper's most pressing recommendations: for curators to tally, identify, and properly house animal remains that may have cultural ties. For archaeologists, this may mean partnering with zoological experts. "Creating a world in which animals in museums are cared for in a culturally connected way requires first caring about identification," says co-author Jimmy Arterberry, a Comanche Nation tribal historian and elder. Then we "can flag items for return or create opportunities for communities with ties to these remains to provide meaningful context for their care."

The report also recommends careful tracing of how animal remains have been treated from the moment of their discovery. That's because animals unearthed during archaeological excavations often get stored with ordinary zoological specimens, severing cultural links that might make them eligible for repatriation, says Kristen Barnett, an Indigenous (Unangax̂) scholar and archaeologist at the University of British Columbia who was not involved in the study. How museums handled those remains also matters, because in past decades, wool textiles, pelts, bones, and feathers were treated with pesticides that may make the objects unsafe to handle today. Today, noninvasive techniques such as near-infrared spectroscopy can reveal traces of chemicals such as arsenic and cyanide. Disclosing any risks is vital to building trust with communities with cultural ties to the remains, the authors say.

Once curators understand something about the history of remains, they can add that information to records for other scholars and descendant communities, says Jane Anderson, a legal scholar and anthropologist at New York University. She co-founded an organization called Local Contexts to help curators and Indigenous communities attach traditional knowledge and biocultural information to digital records. That context can shape how objects are handled and displayed. For example, Cheyenne and Arapaho tribes successfully requested that a Sun Dance buffalo skull housed at the Smithsonian Institution's National Museum of Natural History be stored facing east and upside down to signal that it is ceremonially inactive.

Taylor, who is also the curator of archaeology at the Museum of Natural History at CU Boulder, says the recommendations could extend to remains that have no known connection to human cultures, but may still be important to Indigenous peoples, including remains in natural history or teaching collections. But because the new approach will require staff, time, and money—resources in short supply at museums—he recommends focusing on animal remains with clear links to Indigenous communities. His own experiences show how this can be done.

In 2022, Taylor got state funds to rehouse a collection of 200 bison skulls associated with a mass hunting event near Kit Carson, Colorado, roughly 11,000 years ago. The Lakota have cultural connections to the site, so the museum invited a delegation of Lakota elders to perform a blessing and share suggestions for the skulls' care. The team grouped the herd in storage as it would have been in life, and commissioned new fiberglass molds that better support the skulls while leaving them accessible to handling or viewing. The team also scanned some skulls to create 3D models for Indigenous people who want access for cultural or research purposes but are unable to travel. "Having done some of this engagement work, I can already see all the things I would have done differently, but that uncertainty—or fear over potential costs—is no excuse not to engage," Taylor says.

At the moment, Indigenous communities rely on the good faith of institutions to repatriate animal remains. There have been some success stories, such as the voluntary repatriation of Big Medicine, a white buffalo sacred to the Confederated Salish and Kootenai Tribes. In 2022, the Montana Historical Society returned custodianship of the mounted remains to the tribes upon their request. But some museum professionals want a broader policy solution, such as updating NAGPRA to include animal remains; others urge scientists to take responsibility for returning remains—or perhaps not collect them at all.

For example, archaeologists like Barnett and Sara Gonzalez of the University of Washington aim to leave sites they excavate largely intact. Gonzalez, who studies Indigenous food systems in western North America, sometimes finds animal remains during digs. Rather than collecting material, she maps and studies each item in place, employing low-impact methods such as ground-penetrating radar. "My goal is that when we have completed our research, I'm holding nothing," she says.

Ward is now the NAGPRA compliance officer for the state of Colorado and has broad discretion to determine which of the state's roughly 225,000 artifacts are eligible for repatriation. Many, such as a ceremonial whistle made from an eagle's bone, are outside NAGPRA's purview, but Ward knows they hold special meaning. He hopes that one day, after human ancestors have been fully repatriated, museums will move on to other categories. "Even though these items may not fall under NAGPRA, I have the knowledge and the duty of care, so I pull them aside," he says. ■

Amanda Heidt is a freelance journalist and editor in Moab, Utah.

---

POLICY

# Trump names AI-focused team as leaders of his science office

## Michael Kratsios and Lynne Parker, veterans of Trump's first term, come with technology backgrounds

By **Jeffrey Mervis**

President-elect Donald Trump is bringing back the core of his technology policy team from his previous term to lead the White House Office of Science and Technology Policy (OSTP). The announcement fills the top spot at OSTP much sooner than Trump did last time and likely signals the office will prioritize artificial intelligence (AI).

On 22 December, Trump said he would nominate technologist Michael Kratsios to be director of OSTP and an assistant to the president for science and technology. Computer scientist Lynne Parker was named counselor to Kratsios, a new position, and executive director of the President's Council of Advisors on Science and Technology (PCAST).

For the past 3 years, Kratsios has been managing director of Scale AI, which raises money to help grow AI startups. Parker retired in May after more than 2 decades as a faculty member and senior administrator at the University of Tennessee, Knoxville (UTK).

Trump said the pair will be part of a "brilliant team" led by billionaire David Sacks, his new AI and cryptocurrency czar, who will also chair PCAST. The team will include a White House newcomer, internet entrepreneur and podcaster Sriram Krishnan, who Trump said would be "working closely with Sacks" as OSTP's senior policy adviser on AI.

Kratsios and Parker are known entities to the research community from their previous stints at OSTP and other federal posts. Kratsios was the top political appointee in the office for nearly 2 years until meteorologist Kelvin Droegemeier was confirmed in January 2019 as OSTP director and then served as White House chief technology officer (CTO). Parker, who spent 2 years as a manager in the computing directorate at the National Science Foundation (NSF), led federal AI initiatives for 4 years at OSTP and was deputy CTO before returning to UTK in 2022.

The announcement also calmed fears that Trump might choose an outsider holding controversial views on federal innovation policy to lead the office. "APLU is enthusiastic that President-elect Trump selected two individuals who recognize the importance of science to national competitiveness, health, and economic growth," says Mark Becker, president of the 230-member Association of Public and Land-grant Universities. "OSTP and PCAST play central roles in enhancing science in policymaking," he adds.

Even so, Trump's comments suggest AI could occupy an outsize portion of Kratsios's portfolio in an office that has traditionally overseen the entire $200 billion federal investment in science and technology. They also point to Trump's preference for those with backgrounds in technology over basic research.

Kratsios earned an undergraduate degree in politics from Princeton University in 2008. He was chief of staff to venture capitalist Peter Thiel before coming to OSTP in 2017. By comparison, Droegemeier had spent more than 3 decades as a faculty member and administrator at the University of Oklahoma and served on NSF's National Science Board before joining the Trump administration. Most previous OSTP directors could also boast of a distinguished academic career, although the current head, Arati Prabhakar, has made her mark as a federal science administrator and venture capitalist after earning a Ph.D. in applied physics from the California Institute of Technology.

Kratsios must be confirmed by the Senate as director of OSTP, whose staff of more than 100—many on loan from other federal agencies—belies its tiny, $10 million budget. His role as assistant to the president—a title given to Prabhakar but not Droegemeier—does not require vetting by legislators and affords him direct access to the president. Sacks, whose new job is not subject to Senate confirmation, is expected to work out of a separate White House office and serve as a part-time consultant rather than a full-time federal employee.

In a separate announcement, Trump said he would nominate Emil Michael to lead the Department of Defense's Research and Engineering office, the position Kratsios held for 6 months in the first Trump administration. Michael is a lawyer and serial high-tech entrepreneur. ■

FUNDING

# Historic U.K. medical research units under threat

## New focus on “challenge-led” funding may see decades-old institutes close, scientists warn

*By* **Cathleen O’Grady**

Venerable scientific institutes around the United Kingdom could be forced to close under a new funding scheme, according to an open letter signed by more than 500 scientists that urges Patrick Vallance, the U.K. science minister, to reconsider the changes.

Under the plans, the Medical Research Council (MRC), the U.K.’s primary public funder of medical research, will end its rolling funding of university-hosted research units, many of which have been around for decades. Instead, MRC will award large grants to new Centres of Research Excellence (CoREs), which aim to tackle specific challenges within a 14-year funding period. The first two CoREs, which will focus on gene therapies, were announced on 11 December 2024.

MRC says the changes are necessary to produce “transformative” biomedical research in a fixed time frame.

But the petitioners say the decision was made without transparency or input from the scientific community and will result in the loss of important research infrastructure and institutional memory. The move “seems like, scientifically and strategically, an incredibly backward step,” says Tim Dalgleish, a clinical psychologist at the MRC Cognition and Brain Sciences Unit (CBU) at the University of Cambridge.

Each of the 19 MRC Units—which together receive about £100 million annually from MRC—focuses on a specific research topic, such as biostatistics, prions, or epidemiology. Many have long histories: The MRC Toxicology Unit at Cambridge was founded in 1947 to research industrial hazards, for instance, and the MRC Biostatistics Unit was founded in 1914.

MRC announced the changes to its funding model in July 2022, but many scientists had hoped the new U.K. government elected in July 2024 would reverse the decision. However, Vallance defended the move in the U.K. Parliament in October, precipitating the new open letter.

Although the units can in principle apply to replace their funding with CoRE and other grants, the signatories fear many will shut down. Glasgow’s Social and Public Health Sciences Unit (SPHSU) is already set to close in March after unsuccessfully applying for a CoRE grant, and other units will face a reckoning at the end of their regular 5-year review periods, says William Astle, a biostatistician at the MRC Biostatistics Unit.


The Social and Public Health Sciences Unit at the University of Glasgow will close this year as a result of changes to Medical Research Council funding.

Astle says the CoRE grants can’t fully substitute for existing MRC Unit funding, which pays for research staff, Ph.D. students, technicians, premises, and research equipment. CoRE grants cover fewer staff expenses and do not support a department’s infrastructure, Astle says. They are also smaller—although some MRC Units have had budgets of about £9 million annually, CoRE grants are capped at £3 million.

MRC Executive Chair Patrick Chinnery says that under the new model, units can seek salary support from other sources such as universities and grants. MRC has “worked very closely with universities … so that there can be plans in place to try and help manage this transition,” he adds. The move has received “overwhelming support” across the country and will be a “huge opportunity” for universities that don’t currently host MRC units. The overall level of funding will remain the same under the new system, Chinnery says.

Staff at the units aren’t reassured. “I’m not really convinced that there’s a way to cobble together 25 different grants to replicate what we used to have,” says Michael Anderson, a cognitive scientist at Cambridge’s CBU. The university hosts six MRC units that collectively employ more than 550 staff members.

Others worry basic research will lose out under the new model, which requires CoREs to show their work will have a real-world impact. Many of the topics the units focus on don’t fit a challenge-led and time-limited approach, says Marcia Gibson, who researches social determinants of health at SPHSU: “The ability to take a really long-term perspective on major scientific challenges is just not something you get elsewhere.”

MRC made its decisions behind closed doors and without giving the scientific community the chance to voice concerns, says Jenny Lennox, an official at the University and College Union who helped organize the open letter. Chinnery denies this: The MRC panel that developed the plans included a unit director, he says, and MRC hosted webinars with the universities and unit directors to explain the changes, revising the proposals based on their feedback. But Astle says the process “fell dramatically short of meaningful consultation,” with feedback solicited only after plans were largely finalized. The panel has not published meeting minutes to help the community understand its rationale, Gibson says.

Anderson foresees difficult times for CBU. The uncertainty is likely to trigger an exodus of skilled staff, including the radiographers and physicists needed for brain imaging experiments, he says. “I want to stay if I can. But frankly, you get fed up at a certain point,” he says. “There will reach a point at which this lack of stability will be fatal to a place like this.” ■

PHOTO: JOHN PETER PHOTOGRAPHY/ALAMY

Satellite observations show that a band of tropical clouds has narrowed over 2 decades, allowing more sunlight to hit Earth.

CLIMATE SCIENCE

# Earth's clouds are shrinking, boosting global warming

## Narrowing storm bands may be a surprising and dangerous new feedback of climate change

*By* **Paul Voosen**

For more than 20 years, NASA instruments in space have tracked a growing imbalance in Earth's solar energy budget, with more energy entering than leaving the planet. Much of that imbalance can be pinned on humanity's emissions of greenhouse gases, which trap heat in the atmosphere. But explaining the rest has been a challenge. The loss of reflective ice, exposing darker ground and water that absorb more heat, isn't enough to explain the deficit, and the decline in light-reflecting hazes as countries clean up or close polluting industry falls short as well. "Nobody can get a number that's even close," says George Tselioudis, a climate scientist at NASA's Goddard Institute for Space Studies.

But Tselioudis and his colleagues now think they can explain the growing gap with evidence collected by a remarkably long-lived satellite. They find that the world's reflective cloud cover has shrunk in the past 2 decades by a small but tangible degree, allowing more light in and boosting global warming. "I'm confident it's a missing piece. It's *the* missing piece," says Tselioudis, who presented the work last month at a meeting of the American Geophysical Union.

Climate scientists now need to figure out what's causing these cloud changes. They also need to tackle a more alarming question: whether the trend is a feedback of climate change that might accelerate warming into the future, says Michael Byrne, a climate dynamicist at the University of St. Andrews. Although some models have predicted the cloud changes, Byrne says, "I don't think we can answer this question with much confidence."

Clouds come in all shapes and sizes, but two of the most consistent cloud swaths are formed by Earth's large-scale airflow patterns. One band, near the equator, stretches around the planet like a belt. It forms as trade winds of the Northern and Southern hemispheres converge, forcing moist air upward to cool and condense into clouds. Another band occurs in the midlatitudes, where jet streams usher large swirls of stormy weather around the planet.

In August 2024, Tselioudis and his coauthors reported that over the 35 years covered by weather satellite imagery, the equatorial cloud bands had narrowed, while the tracks of midlatitude storms had shifted toward the poles, hemming in the region in which they can form and shrinking their coverage. But the result, published in *Climate Dynamics*, was stitched together from many different satellites, each with its own quirks and errors, which made the researchers unsure the small trends they detected were real.

Now, the team has turned to a single satellite, NASA's Terra, which has been monitoring the planet for nearly a quarter-century. Looking at the same cloud systems, the team found exactly the same trends, with cloud coverage falling by about 1.5% per decade, Tselioudis says. "It's only now that the signal seems to be coming out of the noise." Bjorn Stevens, a climate scientist at the Max Planck Institute for Meteorology, says a couple of percentage points may not sound important. "But if you calculate these trends, it's massive," he notes. "This would indicate a cloud feedback that's off the charts."

The team also found that 80% of the overall reflectivity changes in these regions resulted from shrinking clouds, rather than darker, less reflective ones that could be caused by a drop in pollution. For Tselioudis, this clearly indicates that changes in atmospheric circulation patterns, not pollution reductions, are driving the trend.

The new work doesn't stand alone. Earlier this year, in *Surveys in Geophysics*, a group led by climate scientist Norman Loeb at NASA's Langley Research Center also traced the gap in the energy imbalance to declining cloud coverage. But Loeb, who leads work on the set of NASA satellite instruments called Clouds and the Earth's Radiant Energy System, which tracks the energy imbalance, thinks pollution declines may be playing an important role in the cloud changes, especially in the Northern Hemisphere. "The observations are telling us something is definitely changing," he says. "But it's a complicated soup of processes."

If global circulation changes are at work, an urgent question is whether they will continue, says Tiffany Shaw, a climate dynamicist at the University of Chicago. The same models that predict a narrowing of the equatorial storm belt also suggest climate change will cause air over the eastern Pacific Ocean to warm faster than the west, weakening an important branch of the large-scale circulation. But for the past few decades, the eastern Pacific has actually been cooling, strengthening these winds instead. Other observations, meanwhile, suggest the rest of the circulation is weakening. The confusion makes it hard to know whether the cloud banks will continue to shrink as the world warms. "With a lot of this," Shaw says, "the real world will show us the answer."

Stevens, for one, is increasingly worried. "If the trend holds up, we're in trouble," Stevens says. "We hope, hope it changes its direction tomorrow." ■

PHOTO: NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

# AFTER ABANDONMENT

## When farmers leave their land, what version of nature takes over?

By **Dan Charles**,
*in Tyurkmen, Bulgaria*

People in this small village didn't know what to think when Gergana Daskalova showed up and moved into her grandparents' former house. Ambitious young people generally flee this sleepy village. Few return for long.

Yet Daskalova, a fellow in the Georg August University of Göttingen's Department of Conservation Biology, is determined to live at least part time in the place where she spent childhood summers. Along with all the fond memories, it's become a research site. She's studying what happens to ecosystems when people vanish.

The population of Tyurkmen has fallen by about 80% over the past century. About half of the houses stand empty, with collapsed roofs and teetering walls. Shrubs and small trees are taking root in former gardens.

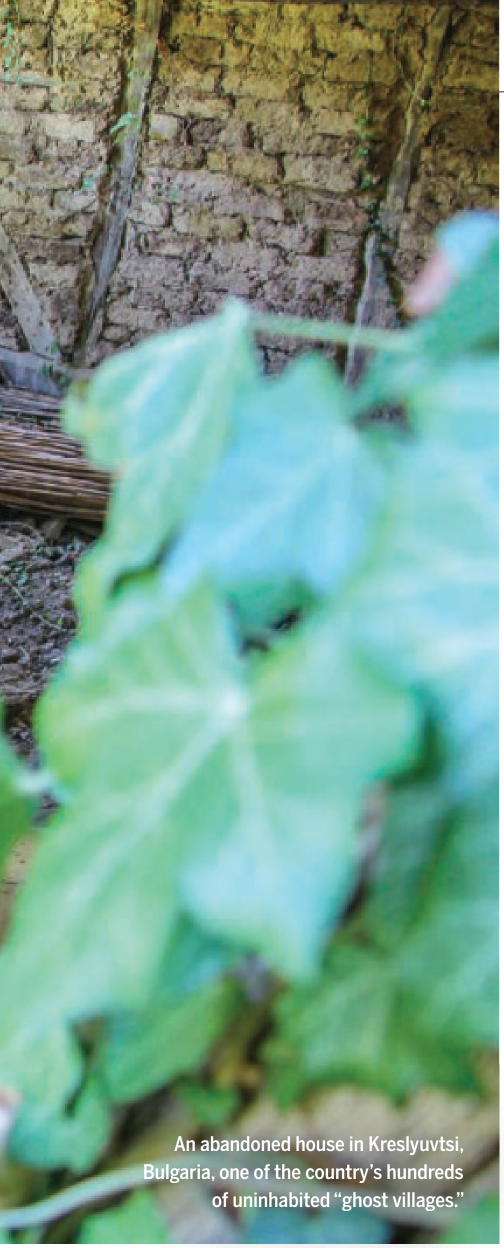
Beyond the village lies open land. On much of it, crops no longer grow, and sheep no longer graze. The plots of land are owned by heirs who live far away. "Every generation, it becomes more complicated to do anything with that land," Daskalova said as she stood on a hill and took in the view in September 2024. "My family has fields that we inherited from my grandparents," she said. "But I don't know where they actually are."

Fallow fields are now covered with opportunistic grasses. Thorny shrubs and wild plum trees are taking root. Stands of an invasive tree called Tree of Heaven are spreading fast.

When Daskalova was growing up in the half-abandoned village, such scenes were simply the backdrop of ordinary life. "I had to leave my village and do my education abroad to realize that, 'Oh, other people don't see this every day,'" she said. The abandonment of rural land "is happening behind the scenes. Most people have no idea that it's occurring."

In reality, it's surprisingly common. "This is a worldwide phenomenon," says Peter Verburg, a land use researcher at the Free

PHOTO: D. CHARLES/*SCIENCE*



An abandoned house in Kreslyuvtsi, Bulgaria, one of the country's hundreds of uninhabited "ghost villages."

University Amsterdam. Global trade in food has fueled the clearing of forests in Brazil and Bolivia for agriculture, but elsewhere it has sidelined small farms with rocky soil, steep hills, or scarce water. "People give up because they cannot compete," Verburg says.

Farmers, or their children, are walking away from land in Eastern Europe, southern France, South Korea, Japan, and mountainous parts of India. It's difficult to measure the exact extent of the trend. Land is often abandoned, then reclaimed and farmed again. But an estimated 120 million hectares have been left fallow in Europe alone since 1990. Globally, the figure since 1950 could be as high as 400 million hectares—half the area of Australia. "Abandonment will continue, I think there's no doubt," Verburg says. In fact, climate change is likely to accelerate it as droughts afflict more farming areas.

The phenomenon raises thorny questions that ecologists and policymakers are now debating. What sort of nature will reclaim this land? Does it add up to environmental restoration or degradation? Should policymakers step in to steer the fate of the land or even stop it from being abandoned?

The answers can be murky. They sometimes depend less on data than on the version of nature that scientists and policymakers value most.

"It's a bit of a cultural war," says Henrique Pereira, an ecologist at the German Centre for Integrative Biodiversity Research and Martin Luther University Halle-Wittenberg. Pereira has led a push in Europe to embrace land abandonment as an opportunity to let nature take its course—even if it opens the way to weeds and invasive species, as in Daskalova's village. "I endorse ecological succession after abandonment," he says. "I'm a 'if you love them, set them free' kind of person." Other ecologists say a hands-off approach is a recipe for loss. When traditional farming disappears, they say, communities of wildflowers and birds that thrive in a mosaic of fields and meadows also vanish. The idea that vibrant ecosystems will naturally emerge instead is "an ecological fairy tale," says Michael Glemnitz, a biodiversity researcher at Germany's Leibniz Centre for Agricultural Landscape Research.

**ON THAT HOT LATE-SUMMER DAY** in September, Daskalova stepped carefully across a rickety footbridge and climbed a grass-covered road through what's left of Kreslyuvtsi. It's one of 30 villages where she and her research team are counting plant species and hiding battery-powered audio recorders in trees to capture the sounds of birds and bats. For each study site that's experiencing abandonment, the researchers have chosen a twin, another village that's similar but is not losing people, in an attempt to isolate the effects of depopulation.

"These changes in villages have always been something very close to my heart," she said. "I really wanted to see, what does human migration mean for nature, for biodiversity? And what are some of the ways that we could reimagine villages in a way that is good for both people and for nature?"

Kreslyuvtsi is officially uninhabited, one of Bulgaria's 500 or so "ghost villages." Like many of them, it's located amid forested mountains.

Daskalova walked past brick and stucco houses with red tile roofs, many of them partially hidden behind a veil of vines. Thickets of wild blackberries filled what once were gardens or orchards. "As far as we can see, it's just brambles," Daskalova said. "The blackberries are suppressing the growth of anything else." A solitary walnut sapling stood amid the shrubs.

The version of nature that's taking over this village, Daskalova said, is a poor substitute for what preceded it. "I think if we were to do a plant survey here right now, we might be at something around like 10 plant species, which for a mountain village is not very many," she said. A village where people still tend gardens and herd goats might have 10 times that many plant species, and also more birds, she said.

There's no sign that time will break the brambles' stranglehold, she said. Another village in this region, abandoned at least 50 years ago, remains so overwhelmed by thorny shrubs that Daskalova couldn't even hack her way through to set up a study site there.

The lack of human activity seems to be choking out biodiversity in mountain villages like this one, where abandonment is most advanced. In the more populous lowlands, on the other hand, where abandoned homes and fields sit alongside others where people still live and work, she has found that some wildlife prospers in the ruins. "Remnants of human infrastructure are actually quite good for biodiversity," Daskalova says. "There will be birds nesting in houses, bats nesting in houses."

But these are just a few data points in a very large and messy array.

**JOHANNES KAMP FIRST** encountered abandoned farmland in 2003, and at the time, it felt like paradise.

Kamp was a 23-year-old bird enthusiast traveling through Kazakhstan, where state-owned agricultural enterprises had collapsed in 1991 along with the Soviet Union. Some 24 million hectares of croplands had been left to their own devices. "These abandoned fields were teeming with wildlife," Kamp says. Foxes patrolled the steppe and the sky "was just absolutely filled with birdsong."

"For me, it was very emotional," says Kamp, who's now a conservation biologist at Göttingen. It seemed a sign of hope in a world where environmental loss is the rule.

In the years that followed, however, grass accumulated on the fallow and ungrazed farmland, and devastating fires erupted. "We had these huge wildfires, sometimes 30,000 hectares," Kamp says. At the same time, new agricultural enterprises showed up to reclaim some of the most productive land. Kamp wondered whether the wildlife paradise he'd witnessed would vanish completely.

Yet the steppe continued to surprise. When Kamp returned to Kazakhstan in the summer of 2024, he found that the danger of fire had receded. Increasing numbers of livestock and a booming population of wild saiga antelopes were keeping the vegetation in check. Meanwhile, the move to reclaim land for farming had stalled.

Abandoned land that's remote or infertile will likely stay that way, Kamp wrote in an email to *Science*.

Today, Kamp can't say whether land abandonment generally is helpful or harmful to biodiversity. "I think there are facets and nuances, but nobody knows what the answer to that question is," he says. "I've changed my mind over the years, to be honest."

Geography and climate certainly affect the answer. Farms in the northeastern United States that went out of business between 1880 and 1930 have turned into mostly healthy forests, with only the occasional stone fence betraying the land's past. (The idea of ecological succession, describing the step-by-step transformation of disturbed land into a vibrant forest, was based on observations of abandoned U.S. farmland.) Yet fallow fields in the arid climate of western Australia can sit almost unchanged for close to a century. "You just have slow erosion, and it's bare ground," Daskalova says.

Scientists in Poland reported in *Science* in 2023 that up to 75% of that country's abandoned farmland is now dominated by invasive species such as goldenrod, walnut, and boxelder maple (21 July 2023, p. 277). Goldenrod, in particular, keeps trees from taking root and reduces populations of wild pollinators and birds, the scientists wrote. Yet a researcher found rising bird populations in abandoned vineyards of southern France. And Kamp has seen "huge populations" of butterflies in a former military training area in Germany that was abandoned nearly 40 years ago—although he expects those populations to decline eventually as open spaces give way to more shrubs and trees.

In Kreslyuvtsi, vegetation is swallowing an abandoned homestead (top) and memorial notices mark the deaths of former inhabitants (bottom).

PHOTOS: (TOP TO BOTTOM) MALKOLM BOOTHROYD; D. CHARLES/*SCIENCE*

In short, the picture is confusing, and researchers are trying to figure out what factors send each landscape down its particular path. There are some prime suspects: whether there's enough rainfall to support a forest, what seeds remain in the soil, and whether diverse ecosystems exist nearby that could colonize the abandoned fields. "It's definitely an opportunity to study," says Emma Ladouceur, a researcher on ecological restoration at the University of Prince Edward Island.

Besides doing fieldwork in villages of Bulgaria, Daskalova is organizing an international working group to collect and analyze ecological data from other abandoned sites around the world. "We want to quantify the different trajectories for what happens," she says.

But the findings won't be able to say how abandoned land should be managed, or whether people should try to manage it at all. "Even if we are scientists, at the end of the day it's going to be a decision based on values," Daskalova says. Those value-based preferences, in fact, are increasingly up for debate.

**EUROPE'S POLICIES** on biodiversity, for example, aim to preserve centuries-old farming landscapes and ecosystems. But Pereira started to question those policies after he returned to Portugal from Stanford University. He'd been inspired by near-wilderness areas in the U.S., and wished for something similar in his home country, "someplace where we have nature working," he says.

In northern Portugal, where he'd carried out earlier fieldwork, he realized that traditional patterns of sheep herding and cattle ranching had collapsed, creating an opening for wilderness to return. He thought: "This is kind of a cool opportunity, you know?"

Without animals to graze the meadows, shrubs and small trees were taking over. Once-rare wild animals, such as roe deer and wild boar, were mounting a comeback.

Pereira took a break from academia in 2006 and served for 3 years as director of Peneda-Gerês National Park, which includes areas long used by ranchers. He decided to change the park's strategy: Instead of fighting the retreat of farming, he'd embrace it. He drafted a new management plan for the park, designating land where animals wouldn't graze any longer.

He soon got a call from Portugal's director of conservation. "You cannot do that," Pereira recalls the official saying, because it would violate the European Union's Habitats Directive. One of Europe's flagship conservation laws, the directive lists more than 200 habitat types that need to be preserved.

Sheep kick up dust as a herder leads them through the streets of Tyurkmen. Traditional farming is declining in many countries, opening the door to land abandonment.

About half, according to Pereira, require some sort of human management.

They include the seminatural meadows of traditional farming communities. Such habitats can be threatened by intensive agriculture, as farmers try to grow more hay, or graze more cattle. Yet if farmers stop all grazing or mowing, they give way to shrubland or forests.

The EU is using its agricultural subsidies as a tool to fight land abandonment, persuading farmers to keep tending such “cultural landscapes.” Landowners who fail to keep their land in “good agricultural condition” can lose their payments. Farms at higher risk of going out of business because they are in remote or mountainous areas get bigger payments.

Pereira found this wrong-headed. He persuaded Portugal’s conservation officials that removing cattle from parts of the park could serve a broader goal of the Habitats Directive—promoting more natural and dynamic ecosystems. (After he stepped down as park director, however, his successor decided not to adopt the proposed changes because of opposition from local cattle ranchers and shepherds.)

Ever since, Pereira has been on a mission to change Europe’s approach to nature conservation. He co-authored a much-cited paper in 2012 arguing that land abandonment could be an opportunity to “rewild” European landscapes.

Preserving static “museum landscapes” should not be a priority, Pereira says. In fact, “species richness” shouldn’t even be the main goal if it depends on human management, as it does at a zoo or botanical garden. Instead, he argues, conservationists should try to create conditions that allow natural ecological process to unfold, free of human control, with shifting communities of plants, herbivores, carnivores, and scavengers such as vultures adapting to disruptions such as climate change.

The problem, though, is that natural succession can take abandoned landscapes in directions that most people find ugly—such as masses of thorny blackberries. Pereira says certain interventions, such as introducing herbivores, might “get more of these landscapes on trajectories that are seen as good for nature and people.”

In addition, researchers are finding that it can take a very long time, perhaps centuries, for complex and dynamic ecosystems to emerge on former farmland. In fact, there’s no guarantee that they ever will.

Pereira says he once expected to see such results within his lifetime. Now, he has come to accept that “there are many things I’m not going to see. Maybe my children, maybe my grandchildren get to see that.”

Pereira hasn’t managed to change Europe’s policies significantly, but his views are getting increasing attention. A group called Rewilding Europe hopes to create wild reserves in 10 landscapes across the continent, from the Scottish highlands to the Danube delta. The EU recently launched a $10 million research project aimed at exploring how rewilding farmland and forests could help Europe reach its biodiversity and climate goals.

Pereira admits that for people with ties to the land, letting nature reclaim it can be hard to stomach. In surveys, they express “a feeling of loss. A feeling of failure,” he says.

But those feelings are sometimes mixed with others, which he’s witnessed firsthand in northern Portugal. “For instance, farmers have this fight against the wolf. You know, ‘The wolf, he’s really an enemy.’ But then a shepherd will say,”—and Pereira’s voice shifted to an awestruck whisper—“The other day I saw a wolf, and oh man, it was quite something.”

**AS DASKALOVA WALKED** through the village where five generations of her family lived, it’s the human losses that seemed to preoccupy her. In the doorway of one ruined house, a child’s book, with a boy’s name written inside, lay open on the ground. “Somebody poured their heart and soul into creating a home, and now it’s all collapsing,” she said. “It’s quite sad. That’s why I find it very difficult when there’s the question of, ‘Should we let abandoned land just be? Should nature claim it back?’ It’s not just birds coming in. It’s little bits of human history disappearing.”

Even the benefits for nature, which she sometimes sees in her data, may be short-lived, Daskalova said, because cheap abandoned land is ripe for exploitation. “Any conservation benefit of restored habitat on abandoned land can disappear in the blink of an eye,” she says.

A mining company recently bought land 500 meters from the edge of Tyurkmen. It wants to blast open a limestone quarry there. Daskalova is organizing a citizens’ movement to stop it, but a village that’s already half-deserted lacks political clout. “The places that are experiencing a lot of abandonment are very vulnerable,” she says.

Some researchers have suggested geographic targeting of EU subsidies, with funding to stop land abandonment where the consequences seem particularly damaging, while promoting rewilding in other areas. But deciding where to intervene would be daunting. Daskalova, for her part, is taking a very personal stand for villages. She’s currently building a new house in Tyurkmen for herself, her husband, and their young son. ■

PHOTO: MALKOLM BOOTHROYD

# INSIGHTS

## PERSPECTIVES

PHOTO: LAKOTA BURWELL AND TASHA SHIELDS

ECOLOGY

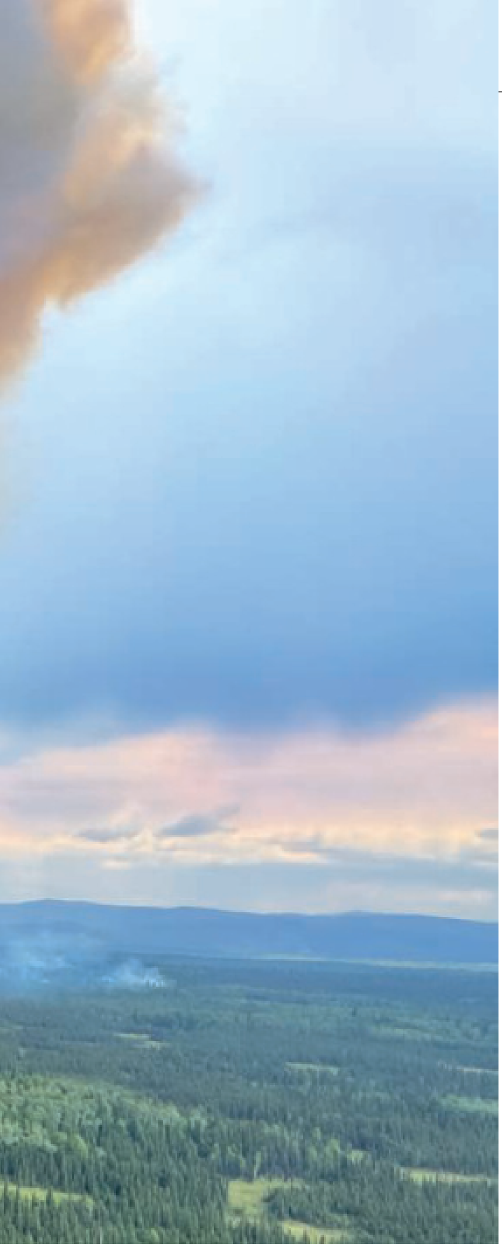
# Disentangling the drivers of wildfires

The risk of wildfires varies across regions with different vegetation

*By* **Jianbang Gan**

Wildfire occurrence and scale worldwide have risen over recent decades, with the most destructive wildfires in North America taking place in the past decade (*1*, *2*). Alleviating wildfire effects requires in-depth understanding of what propels them, which informs better assessment of risks and consequential societal impacts. On page 91 of this issue, Wang *et al.* (*3*) report the key drivers of burn severity—a measure of fire-induced change in vegetation and soil—by building a model using daily fire growth data collected in the past two decades. This analysis reveals the importance of regional characteristics on the ecological effects of intense fires, which corroborates recent conclusions by Jones *et al.* (*4*) on global forest fires and by Balch *et al.* (*5*) on fast-spreading wildfires in the US. Together, these studies identify different drivers of wildfire effects at the national and global levels that could help develop region-specific risk maps for wildfires.

The ignition and behavior of a fire and its effects on a burned area, including the alternation of soil and vegetation and damages to infrastructures, are two important aspects of assessing the threats of wildfires. This information is used to recommend a risk management plan to minimize the impact of a wildfire. Previous efforts have characterized the activity and effect of wildfires as well as their drivers on the basis of historical fire data on both national and global scales (*1*, *6*, *7*). These studies led to the development of predictive models that are used to project wildfire effects under different scenarios over time in locations where no historical fires have been



A wildfire that started near Salcha, Alaska, in June 2024 burned more than 140,000 acres of land.

recorded. However, the drivers of wildfire effects are complex and involve multiple interlinked factors, such as climate, vegetation, topography, and human activity.

Wang *et al.* examined potential drivers of wildfire effects in Canada by grouping four to five individual predictors that measure different attributes into three categories—weather, topography, and fuel. For example, the fuel category included indicators for vegetation type, biomass, crown closure (blocking of sunlight from reaching the forest ground), tree height, and coniferous species coverage. Each indicator showed varying degrees of association with burn severity. Among all the factors, fuel aridity, which reflects the abundance and moisture content of flammable vegetative fuels, was determined to be the primary driver for most Canadian forest fires between 1981 and 2020. The model also showed variation in the effect of individual drivers across different regions of Canada. Wang *et al.* found a large increase in burn severity in northern Canada compared with other regions in Canada. This trend was more notable in years after 2001. Whereas weather condition was the dominant driver of the effect of wildfires in northern Canada, fuel aridity and vegetation type were key drivers of wildfires in southern areas.

Jones *et al.* (*4*) likewise reported that vegetation dynamics (plant growth affected by weather and soil condition) coupled to weather were the primary drivers of forest fire carbon emissions from 2001 to 2023 in the nontropical regions of the world. This is in contrast to their identification of human activities—such as deforestation and land fragmentation—as the primary drivers of fire carbon emission in the tropical and subtropical regions. Jones *et al.* (*4*) also identified nine predictors of forest fires, including soil moisture, forest continuity, terrain, and human population density. These indicators were correlated with burned areas and overall carbon emissions in different regions around the world. These findings indicate area-specific drivers of wildfire effects.

Balch *et al.* (*5*) further highlighted the need for regional models for wildfires rather than relying on coarse climate data. Their study reported fire growth rate as a key contributor to the destructive character of a fire. Fast fires (burning more than 1620 ha per day) that occurred in the US in the past two decades only accounted for 2.7% of total fire incidents, but they caused 78% of the damage to infrastructure. The study also showed that weather and vegetation type dictated the fire growth rate, although human factors, such as arson and downed power lines, may also have contributed to fast fires. A single factor could also cause various effects over time. For example, an increase in precipitation or atmospheric moisture decreased immediate fire risk but could stimulate vegetation growth and accumulate fuel that, if ignited, could yield a more severe fire in the future.

These findings point to different ways of alleviating wildfire effects in both wildlands and residential areas. They show that wildfire management must reflect differences in regional drivers as well as the change in fire effects over time. Developing region-specific strategies that monitor primary indicators and provide mitigation efforts could improve wildfire risk management. Further, wildfire effects could be minimized by altering the conditions of admissible drivers. Although slowing down climate change is important in the long term, treating vegetative fuels by prescribed burns and forest thinning (partial fuel removal) could decrease fuel availability, alter forest structure, and change the composition of vegetation. These approaches may not only suppress ignition but also reduce intensity and speed of fire to subdue detrimental effects (*8*, *9*). Proactively managing grasslands or establishing fuel breaks (a strip of land with vegetation removal) in areas adjacent to residential communities can also reduce wildfire threats by slowing the pace of fires (*10*).

From an ecological perspective, the increase in fire activity in boreal forests, especially in the northern regions of the world, has raised grave concerns about the health and function of biomes that act as important carbon sinks (*11*). These remote areas support complex and delicate relationships between different ecosystem components, such as carbon sinks and species (*11*). Such associations make fuel management more challenging. Cooperation between the US, Canada, and Russia, which share 93% of the global boreal forest, is needed to effectively manage fire while preserving this valuable ecosystem of the northern hemisphere (*12*).

Continuous monitoring of activity, effects, and drivers of wildfire can provide early warning. Expanding the capacity for rapid integration, processing, and sharing of wildfire data in different regions from various sources could inform more-accurate model development and guide wildfire management decision-making. Extending analyses similar to those of Wang *et al.* to other countries is an urgent matter not only to human health and infrastructure but also to vulnerable biomes in the world. ■

---

Department of Ecology and Conservation Biology, Texas A&M University, College Station, TX, USA. Email: j-gan@tamu.edu

## REFERENCES AND NOTES

1. P. Jain *et al.*, *Nat. Commun.* **15**, 6764 (2024).
2. "Unprecedented Scale: Exploring the Largest Wildfires in US History," *Earth.org*, 2 March 2024; https://earth.org/worst-wildfires-in-us-history/.
3. W. Wang *et al.*, *Science* **387**, 91 (2025).
4. M. W. Jones *et al.*, *Science* **386**, eadl5889 (2024).
5. J. K. Balch *et al.*, *Science* **386**, 425 (2024).
6. A. Duane, M. Castellnou, L. Brotons, *Clim. Change* **165**, 43 (2021).
7. M. W. Jones *et al.*, *Rev. Geophys.* **60**, e2020RG000726 (2022).
8. K. T. Davis *et al.*, *For. Ecol. Manage.* **561**, 121885 (2024).
9. S. J. Prichard *et al.*, *Ecol. Appl.* **31**, e02433 (2021).
10. A. B. Massada, V. C. Radeloff, S. I. Stewart, *Int. J. Wildland Fire* **20**, 59 (2011).
11. S. Gauthier, P. Bernier, T. Kuuluvainen, A. Z. Shvidenko, D. G. Schepaschenko, *Science* **349**, 819 (2015).
12. Food and Agriculture Organization of the United Nations (FAO), "Global Forest Resources Assessment 2020" (FAO, 2020).

10.1126/science.adu5463

THERMAL MANAGEMENT

# Compressed ionic plastic crystals are cool

## A family of materials exhibits a large thermal response at subambient temperatures

*By* **Josep-Lluís Tamarit** *and* **Pol Lloveras**

Traditional vapor compressors for residential and industrial cooling create a vicious cycle: Cooling contributes to global warming, which then demands more cooling. These systems are a major source of greenhouse gas emissions because of their limited efficiency and direct leaks of harmful hydrofluorocarbons from billions of appliances worldwide (*1*). Solid-state cooling systems are emerging alternatives that provide more sustainable solutions. They use either the thermoelectric effect to convert electric energy to temperature difference or the caloric effect, which allows reversible temperature changes in response to external stimuli. However, existing solid-state cooling systems are limited by the lack of proper materials. On page 56 of this issue, Piper *et al.* (*2*) describe a family of organic ionic plastic crystals that exhibit a pressure-driven caloric effect. This potentially large group of materials could provide energy-efficient cooling at low temperatures suitable for residential and industrial applications.

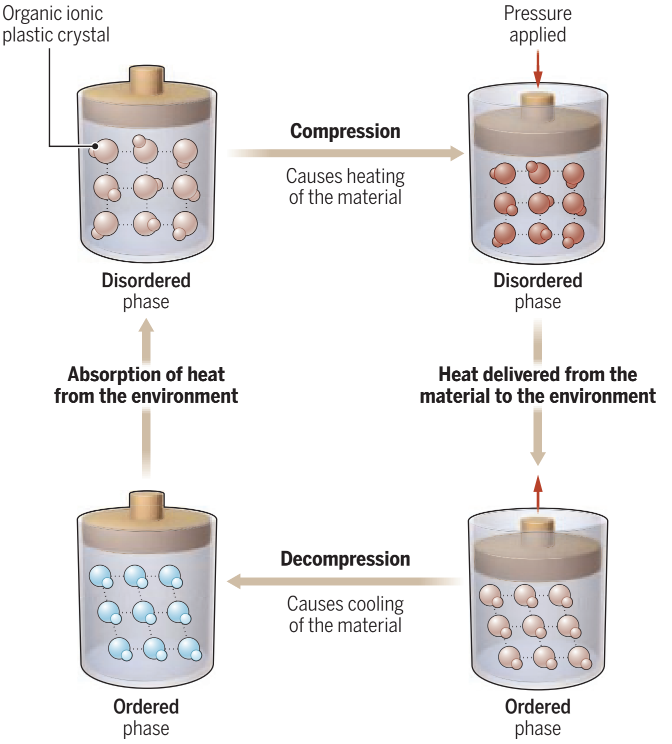
Caloric effects in solids can provide energy-efficient cooling without releasing hazardous gases to the atmosphere. However, a caloric material often requires a large external field (magnetic, electric, or mechanical) to induce a reversible thermal response. Among various caloric effects, the barocaloric effect—pressure-driven heat adsorption or release—has demonstrated the level of cooling that is close to that of fluid refrigerants in the vapor compressor systems (*3*, *4*). Over the past decade, the required pressure to induce a temperature change decreased by half while achieving a 10-fold increase in the refrigerant capacity compared to Heusler alloys, which are magnetic intermetallic materials (*5*). This outperforms refrigeration from magnetocaloric and electrocaloric effects (*6*).

One type of barocaloric material is a crystalline solid that absorbs or releases heat upon an order-disorder phase transition. A so-called colossal caloric effect—a substantial disorder during a phase transition that causes a large entropy change—that was observed during the order-disorder phase transition of an organic plastic crystal points to the potential of the barocaloric effect for solid-state cooling systems (*3*, *4*). Plastic crystals are formed by weakly interacting molecular entities that are orientationally disordered. They can transform reversibly between ordered and disordered (plastic) crystalline phases. An isothermal entropy change, which is a figure of merit for evaluating the performance of a caloric material, of a plastic crystal exceeds more than 100 J kg$^{-1}$ K$^{-1}$ (*3*). This is 10-fold greater than that of other caloric materials. A similar colossal caloric effect has been reported in other solids, including organic and inorganic plastic crystals (*3*, *4*), ferroelectric plastic crystals (*7*), inorganic ionic plastic crystals (*8*), perovskite-structure oxide (*9*), dialkylammonium metal-halides (*10*), dialkylammonium salts (*11*), and spin-crossover compounds (*12*). However, existing colossal caloric materials have shown order-disorder phase transition temperatures that are not within a range suitable for cooling applications.

> **"This would increase the possibilities of finding better candidates for a solid-state cooling effect."**

### Pressure-driven solid-state cooling

A family of organic ionic plastic crystals undergo an order-disorder transition under an applied pressure at subambient temperatures. This solid-state phase transition can provide cooling at operation temperatures that are relevant to many residential and industrial applications.



GRAPHIC: A. MASTIN/*SCIENCE*

Piper *et al.* report a family of solids that exhibit a colossal caloric effect, called organic ionic plastic crystals. Although this group of materials has long been known for solid-state electrolytes (*13*–*15*), their barocaloric effect has not been studied. An organic ionic plastic crystal commonly consists of two small or globular organic molecules (a cationic and an anionic) linked by weak ionic forces. The material undergoes a disorder-to-order solid phase transition when an external pressure is applied (see the figure), releasing heat to the surrounding environment. By contrast, when the applied pressure is released, the material transitions into an ordered phase and subsequently absorbs heat from the surrounding environment to provide a pressure-driven cooling effect.

Piper *et al.* characterized four prototype organic ionic plastic crystals and identified

Grup de Caracterització de Materials, Departament de Física and Barcelona Research Center in Multiscale Science and Engineering, Universitat Politècnica de Catalunya, EEBE, Campus Diagonal-Besòs, Barcelona, Catalonia, Spain. Email: josep.lluis.tamarit@upc.edu

two materials that exhibit substantially high entropy changes. The results revealed two functional properties: ionic diffusion in the disordered phase, which could be useful for potential solid electrolyte application, and a highly energetic nonisochoric first-order phase transition (with a volume change) between the ordered and disordered phases, at which colossal barocaloric effects can be harvested. The reported organic plastic crystals not only maintained the desired properties of the plastic crystal's predecessors but also displayed additional improved properties. It showed an enhanced pressure-driven transition temperature shift with a temperature range of operation appropriate for mass-market applications that is below the cooling point but close to ambient. Piper *et al.* also added alumina to the organic ionic plastic crystals to further decrease hysteresis (temperature difference between onset heating and onset cooling). The small hysteresis reduced the minimum applied pressure required for the reversible phase transition. This decrease in the required pressure could provide safe and fast heat transfers with reduced cost, weight, and size of a cooling device.

Although the organic ionic plastic crystals reported by Piper *et al.* are promising for solid-state cooling systems, further assessment of their chemical flexibility is needed to expand the materials library. This would increase the possibilities of finding better candidates for a solid-state cooling effect. Beyond cooling, materials that display the colossal caloric effect could be an energy-efficient replacement for traditional fuel or electric burners. Organic ionic plastic crystals thus open a pathway to sustainable cooling and heating systems. ■

### REFERENCES AND NOTES

1. A. Morlet *et al.*, The Impact of the Refrigeration Sector on Climate Change, 35th Informatory Note on Refrigeration Technologies (International Institute of Refrigeration, 2017).
2. S. L. Piper *et al.*, *Science* **387**, 56 (2024).
3. B. Li *et al.*, *Nature* **567**, 506 (2019).
4. P. Lloveras, Ed., *Barocaloric Effects in the Solid State: Materials and Methods* (IOP Publishing, 2023).
5. L. Mañosa *et al.*, *Nat. Mater.* **9**, 478 (2010).
6. X. Moya, N. D. Mathur, *Science* **370**, 797 (2020).
7. A. Salvatori *et al.*, *J. Mater. Chem. A Mater. Energy Sustain.* **11**, 12140 (2023).
8. M. Zeng *et al.*, *Adv. Sci.* **11**, e2306488 (2024).
9. Y. Kosugi *et al.*, *Adv. Funct. Mater.* **31**, 2009476 (2021).
10. J. Li *et al.*, *Adv. Funct. Mater.* **31**, 2105154 (2021).
11. J. Seo *et al.*, *J. Am. Chem. Soc.* **146**, 2736 (2024).
12. M. Seredyuk *et al.*, *Adv. Funct. Mater.* **34**, 2315487 (2024).
13. D. R. MacFarlane *et al.*, *Nature* **402**, 792 (1999).
14. D. R. MacFarlane, M. Forsyth, *Adv. Mater.* **13**, 957 (2001).
15. J. M. Pringle *et al.*, *J. Mater. Chem.* **20**, 2056 (2010).

10.1126/science.adu3670

Female noctules expend less energy when they migrate during warm nights with favorable tailwinds.

PHOTO: PAUL HOBSON/NPL/MINDEN PICTURES

ECOLOGY

# Bat tracking across Europe

## New technology answers questions about the seasonal migration of bats

*By* **Liam P. McGuire**

They are small, only come out at night, are silent (to humans), and can fly hundreds of kilometers a night. These factors make studying migrating bats particularly difficult. But overcoming these challenges opens opportunities to better understand migration, a widespread phenomenon in the animal kingdom. Comparative studies across taxa can offer insights into general rules that govern how animals interact with their environments, and species-specific studies can provide key information for developing effective conservation strategies. On page 97 in this issue of *Science*, Hurme *et al.* (*1*) report the use of new Internet of Things technology to investigate movement patterns of migratory bats, the influence of weather, and trade-offs that bats face in decisions about the timing and energetic cost of long-distance flights.

Compared to other migrating animals, little is known about bat migration, even basic information about where some species go during migration. Outstanding mysteries include the factors that influence decisions about where and when to travel and where to stop en route, how bats cope with conditions in the aerosphere, and the physiological mechanisms that underpin their long-distance journeys. Addressing these questions requires consideration of variation within and among species. Comparisons across species and taxonomic groups may reveal common strategies and the evolutionary drivers and fitness consequences of migration.

Some of the earliest reports of bat migration date to the late 1800s (*2*). Yet, more than a century later, the movements of migrating bats are still difficult to determine (*3*). Early studies focused on mark-recapture methods (catch an animal, mark it, then hopefully recapture it somewhere else), but this approach is generally limited by low recapture rates (*4*). The advent of biotelemetry presented exciting new opportunities to follow the movements of individual bats by attaching a small radiotransmitter to the animals and locating the signal (typically over relatively small distances) with a receiving unit (*5*). However, the small body size and the spatial scale at which migrating bats traverse the landscape limit the possibilities of documenting migration (transmitters must be small with a short battery life, and bats can quickly fly out of range of the receiver). A relatively small number of large-bodied bats can carry a device that uses Global Positioning System (GPS) satellites to track their movements (*6, 7*). Although smaller GPS units are now available, they cannot transmit signals and must be recovered from bats. Thus, although such devices provide fascinating data, the sample size is limited (*8*). Using aircraft to track the sig-

---

Department of Biology, University of Waterloo, Waterloo, ON, Canada. Email: liam.mcguire@uwaterloo.ca

nals from bats carrying radiotransmitters has been informative (*9, 10*) but presents logistical and financial challenges. Automated radiotelemetry networks, such as the Motus Wildlife Tracking System, use a network of receiving stations positioned around the world to detect signals from radio-tagged animals (*11*). This approach has been very effective for many species of migrating animals and shows great potential for documenting long-distance movements of migrating bats.

Hurme *et al.* have overcome the challenges of tracking migrating bats with a new technology based on the Internet of Things—that is, a method that relies on a network of stations that receive information from tags on bats without human intervention. The authors developed a new biotelemetry tag that not only records location information but also measures physical activity with accelerometry, as well as temperature of the surrounding environment. Notably, the bat does not need to be recaptured, and the tag does not need to be recovered. The device records information every 60 s and transmits summary data to the network once per day, providing location, daily maximum and minimum temperature, proportion of active time (from accelerometry), and an index of daily energy expenditure (total vectorial dynamic body acceleration). Hurme *et al.* tracked migrating common noctules (*Nyctalus noctula*), which migrate up to ~1500 km across Europe. As for many hibernating bats, common noctules mate before hibernation and females become pregnant in the spring. Migration is female-biased, with pregnant females migrating from winter hibernation sites to summer maternity areas. The authors collected up to 46 days of tracking data for females that migrated over distances of up to 1116 km, with single night flights of up to 383 km, much greater distances than previously documented.

Hurme *et al.* took their tracking study well beyond a basic description of movement patterns. By integrating multiple types of data, the authors gleaned new insights into the drivers of migration and the basis for migration decisions. For example, bats can select nights with weather conditions that are advantageous for migrating. Specifically, noctules migrated primarily on relatively warm nights with favorable tailwinds and lower crosswind speeds. Such conditions allowed bats to fly at lower air speeds, which required less energy expenditure. Furthermore, flexibility in the timing of migration was associated with energetic trade-offs. In the first half of the spring migration season (mid-April to early May), bats were more likely to journey on nights with better wind support. Later in the season (mid-May to early June), pregnant females face increasing energy costs of flight because of the increasing weight of the developing embryo. Thus, there is a trade-off between waiting for favorable flight conditions and the risk of increased flight costs. Indeed, Hurme *et al.* observed that bats that migrated later in the season experienced less supportive wind conditions and greater energetic costs.

The findings of Hurme *et al.* provide excellent data to inform many pressing questions about bat migration, including the influence of weather conditions, and decisions the animals make regarding the costs and risks of flight. Although the study highlights that there is much that is still not well understood about migrating bats, it is important not to discount what is known. Populations of many migratory bat species around the globe are increasingly imperiled *(12)*, primarily because of wind turbines that kill hundreds of thousands of migrating bats each year (*13*). However, restricting turbine operation in conditions that place bats at risk reduces bat mortality (*14*), and therefore it is possible for green, renewable wind energy and sustainable bat populations to coexist. Bats are important components of ecosystems and offer exciting opportunities for comparative studies to understand evolutionary patterns and drivers of migration and the interactions between animals and their environments. Studies that leverage new technologies or approaches can reveal previously unknown aspects of these understudied animals, from ecology to behavior, physiology, and evolution. But if action is not taken to address threats facing bat populations, they may not be around much longer to study. ■

## REFERENCES AND NOTES

1. E. Hurme, I. Lenzi, M. Wikelski, T. Wild, D. K. N. Dechmann, *Science* **387**, 97 (2025).
2. C. H. Merriam, *Trans. R. Soc. Can.* **4**, 85 (1887).
3. P. M. Cryan, R. H. Diehl, "Analyzing bat migration" in *Ecological and Behavioral Methods for the Study of Bats*, T. H. Kunz, S. Parsons, Eds. (Johns Hopkins Univ. Press, ed. 2, 2009), pp. 476.
4. D. R. Griffin, "Migration and homing of bats" in *Biology of Bats*, vol. 1, W. A. Wimsatt, Ed. (Academic Press, 1970), pp. 233.
5. J. Clerc, R. M. Brigham, J. G. Boyles, L. P. McGuire, "A NASBR history of radiotelemetry: How technology has contributed to advances in bat biology" in *50 Years of Bat Research*, B. K. Lim *et al.*, Eds. (Springer, 2021), pp. 241.
6. N. Sapir, N. Horvitz, D. K. N. Dechmann, J. Fahr, M. Wikelski, *Proc. Biol. Sci.* **281**, 20140018 (2014).
7. H. V. Richter, G. S. Cumming, *J. Zool.* **275**, 172 (2008).
8. T. J. Weller *et al.*, *Sci. Rep.* **6**, 34585 (2016).
9. D. K. N. Dechmann *et al.*, *PLOS ONE* **9**, e114810 (2014).
10. M. T. O'Mara, M. Wikelski, B. Kranstauber, D. K. N. Dechmann, *Ecology* **100**, e02762 (2019).
11. P. D. Taylor *et al.*, *Avian Conserv. Ecol.* **12**, 8 (2017).
12. W. F. Frick *et al.*, *Biol. Conserv.* **209**, 172 (2017).
13. C. C. Voigt *et al.*, *Bioscience* **74**, 240 (2024).
14. M. D. Whitby, M. T. O'Mara, C. D. Hein, M. Huso, W. F. Frick, *Ecol. Solut. Evid.* **5**, e12371 (2024).

10.1126/science.adu7970

CANCER

# Inherited genome instability

## Germline structural variants are a risk factor for pediatric extracranial solid tumors

*By* **Jayne Y. Hehir-Kwa**[1] *and* **Geoff Macintyre**[2]

Extracranial solid tumors, such as neuroblastoma, Ewing sarcoma, and osteosarcoma, are a leading cause of morbidity and mortality in children. In contrast to adult cancers, in which exogenous mutagens or age-accumulated DNA damage can drive tumor development (*1*), childhood cancers lack the extended time frame needed to amass the mutations required for tumorigenesis by those routes. Therefore, endogenous mutagenic processes are a likely source for cancer-enabling mutations in pediatric cancers. Yet the detailed biological processes leading to tumorigenesis in these cancers remain mostly unknown. On page 39 of this issue, Gillani *et al.* (*2*) report that rare germline structural variants (SVs), which are a family of large DNA rearrangements that vary in size from 50 to millions of nucleotides, are a risk factor for pediatric extracranial solid tumors. The findings suggest that germline SVs may contribute to early genome instability in these cancers and could inform the design of targeted therapies.

Pediatric cancers have a higher inherited predisposition than adult cancers, but the underlying factors that lead to the initiation of tumorigenesis in children remain unclear (*3*). Previous studies have mainly analyzed small germline variants in childhood cancers, such as single-nucleotide variants or small insertions or deletions (indels), which has led to the identification of risk variants in cancer-related genes (*4*, *5*). Rather than revealing alterations in oncogenes, most of these small variants appear to target genes involved in maintaining genome fidelity, including disruption of the gene *MutS homolog 2* (*MSH2*), which can result in the accumulation of single-base substitutions or small indels, and alterations of genes that can result in structural DNA damage accumulation [includ-

---

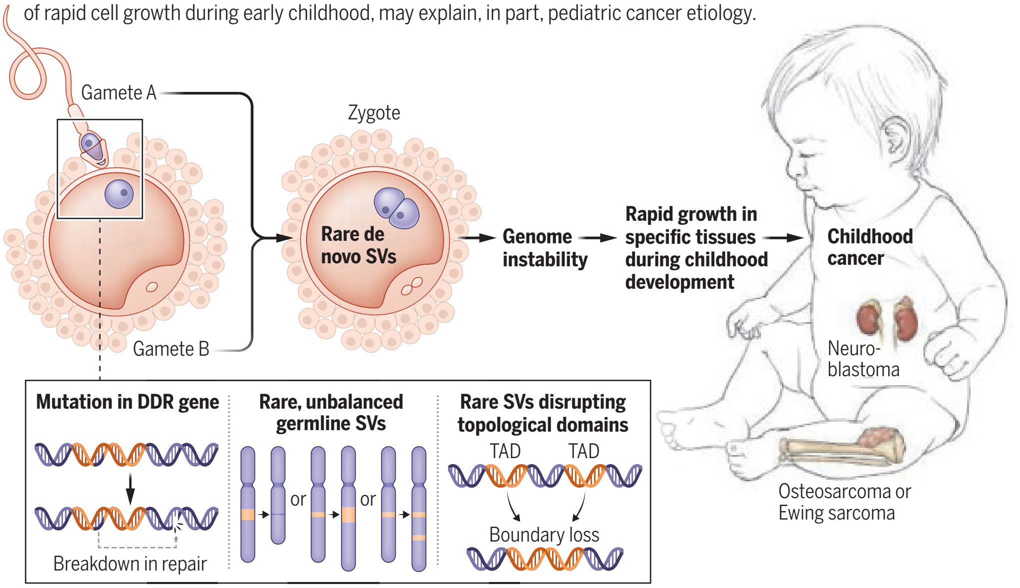
[1] Princess Máxima Center for Pediatric Oncology, Utrecht, Netherlands. [2] Spanish National Cancer Research Center (CNIO), Madrid, Spain. Email: gmacintyre@cnio.es

ing the genes *tumor suppressor protein 53* (*TP53*), *BRCA1-associated RING domain gene 1* (*BARD1*), *protection of telomeres 1* (*POT1*), and *alpha-thalassemia/mental retardation syndrome X–linked* (*ATRX*)]. Beyond these small variants, there is an increasing recognition that germline SVs can also contribute to pediatric cancer risk (*6*), but how they promote oncogenesis remains understudied.

Gillani *et al.* analyzed a collection of rare germline SVs associated with childhood cancer predisposition from 1765 affected children. They also analyzed 943 unaffected parents of these patients to ascertain whether the mutations were inherited or acquired de novo. The study also included 6665 adult controls, unrelated to the children.

### Germline structural variants in pediatric tumors

Inherited and de novo acquired, rare structural variants can contribute to increased risk of childhood extracranial solid cancers by causing genome instability. This genome instability, coupled with periods of rapid cell growth during early childhood, may explain, in part, pediatric cancer etiology.



DDR, DNA damage repair; SVs, structural variants; TAD, topologically associating domain.

Germline genome sequencing revealed 84 rare, large (>1 Mb) unbalanced chromosomal abnormalities (genetic material is added or lost) that increased the risk of solid tumors in children, especially in males. These included SVs inherited from an unaffected parent (82%) and de novo–acquired (18%) abnormalities. Most of these large SVs did not overlap with cancer-related genes or known pathogenic loci. This is somewhat expected because large SVs that disrupt genes tend to cause severe phenotypes and are thus rare in the population (*7*). Gillani *et al.* propose that the DNA dosage imbalance caused by large-scale aberrations may promote early genome destabilization, alongside a subset that may contain yet unknown pathogenic loci.

Other types of germline SVs were also risk factors for pediatric tumors. For example, rare gene-disruptive germline SVs, which are smaller but more abundant than large SVs, were in pediatric solid tumors but not in samples from control individuals. In neuroblastoma and Ewing sarcoma, these included SVs that may cause loss of function of DNA damage repair genes such as *BARD1*, as well as SVs that affect genes involved in tumorigenesis. Noncoding SVs disrupted gene regulation by perturbing topologically associating domain (TAD) boundaries, particularly in neuroblastoma. Not only did this cause dysregulation of genes residing within these TADs, but it may also promote genome instability through possible collisions between the replication and transcription machinery (*8*). Collectively, these findings suggest a model whereby germline variants contribute to cancer in children by disrupting genome stability (see the figure). This is mirrored in hereditary disorders such as xeroderma pigmentosum, Li-Fraumeni syndrome, Fanconi anemia, ataxia-Telangiectasia, and constitutional mismatch repair deficiency, which all show increased cancer incidence during childhood and are characterized by variants that promote genome instability (*9*).

Large-scale studies of the somatic changes found in the genomes of childhood cancers suggest a lower mutational burden in children compared with adults (*10*). However, this may be explained in part by the opportunity for adult cancers to have long-term accumulation of passenger variants (mutations that do not directly drive cancer initiation and progression) through age-related mutational processes (*11*). Thus, a similar average burden of cancer-causing mutations might be required for cancer initiation in both adult and childhood cancers (*12*, *13*). However, the shorter period of tumor development in children compared with adults requires the rate of mutation to be higher in childhood cancers. This increase would also likely need to occur before the cell has acquired neoplastic potential. A possible mechanism to accelerate the acquisition of mutations in cells already engendered by genome instability could involve an increase in division rate (*14*). This can be readily achieved in childhood during critical periods of growth during which cells naturally proliferate faster. The hijacking of existing cellular processes is supported by the findings of Gillani *et al.*, showing that single gene-disruptive germline SVs tend to target gene pathways related to neurogenesis. As well, neural developmental disorders converge on pathways that are also affected in cancer (*15*) and that there is a tendency for pediatric cancers to originate in precursor cells of non–self-renewing tissues.

Because genome instability appears to be a major rate-limiting factor in the development of pediatric cancers, there may be an opportunity to consider the use of therapies that exploit defects in the cell's DNA damage response. Thus, inhibitors of proteins such as ataxia telangiectasia and rad3–related protein (ATR), wee1 kinase (WEE1), checkpoint kinase 1 (CHK1), or poly(ADP-ribose) polymerase 1 (PARP) could be valuable in this context. In addition, the efficacy of platinum-based chemotherapy, which are platinum-based drugs that bind to DNA and induce damage and cell death, supports the rationale of therapeutically exploiting the DNA damage found in these cancers. These strategies could complement the growing use of immunotherapy in childhood cancers, offering a powerful combination to tackle these aggressive malignancies. Adopting a more precise, targeted approach against pediatric extracranial solid tumors may not only enhance treatment outcomes but could also reduce the late effects seen in pediatric patients that result from long-term toxicities as well as lower the risk of secondary malignancies associated with traditional therapies. ■

## REFERENCES AND NOTES

1. V. L. Cannataro *et al.*, *Mol. Biol. Evol.* **39**, msac084 (2022).
2. R. Gillani *et al.*, *Science* **387**, 39 (2025).
3. J. Zhang *et al.*, *N. Engl. J. Med.* **373**, 2336 (2015).
4. J. Kim *et al.*, *J. Natl. Cancer Inst.* **116**, 149 (2024).
5. L. Mirabello *et al.*, *JAMA Oncol.* **6**, 724 (2020).
6. J. Kim *et al.*, *JCO Precis. Oncol.* **5** (2021).
7. E. E. Eichler, *N. Engl. J. Med.* **381**, 64 (2019).
8. L. Goehring *et al.*, *Annu. Rev. Genet.* **57**, 157 (2023).
9. R. Sharma *et al.*, *Front. Pediatr.* **8**, 570084 (2020).
10. S. N. Gröbner *et al.*, *Nature* **555**, 321 (2018).
11. M. Gerstung *et al.*, *Nature* **578**, 122 (2020).
12. S. W. Brady *et al.*, *Nat. Genet.* **54**, 1376 (2022).
13. I. Martincorena *et al.*, *Cell* **171**, 1029 (2017).
14. N. T. Umbreit *et al.*, *Science* **368**, eaba0712 (2020).
15. B. R. Yavuz *et al.*, *NPJ Genom. Med.* **8**, 37 (2023).

10.1126/science.adu6125

GRAPHIC: N. BURGESS/*SCIENCE*

POLICY FORUM

CLIMATE POLICY

# Solar geoengineering research faces geopolitical deadlock

## The permissive “science first” approach has failed, but a nondeployment deal might yet enable responsible research

By **Duncan McLaren**[1] and **Olaf Corry**[2]

As a powerful potential lever on global temperatures, solar geoengineering—or solar radiation modification (SRM)—comes with formidable environmental and political risks. Assessments of SRM acknowledge that future governance of “direction, ownership and control” [(*1*), p. 45] is no optional extra: Outcomes achieved in modeling studies presuppose governance “to restrain harmful or unjust use of SRM…and assess and minimize any countervailing harms” [(*2*), p. 6]. But, although research is deemed urgent, the “guardrails” and institutions needed remain moot (*3*) or are deferred for later (*4*). Tentative efforts toward international cooperation and oversight collapsed most recently in February 2024. Drawing upon observations of the United Nations Environment Assembly (UNEA), we set out how clashing geopolitical interests and diverging knowledge politics are making even research controversial. Although the differences seem irreconcilable, we suggest that apparent consensus around nondeployment of geoengineering could be seized upon to provide some hope for coordinated and responsible SRM research.

### GOVERNANCE DEFERRED?

Scientific assessments of SRM by institutions such as the Royal Society and the National Academies recognize the need for global governance to precede any deployment of SRM but tend to presume that more SRM research can safely be pursued in the meantime, especially if it conforms to some form of ethical guidance. More scientific research, some suggest, would reduce uncertainties and thereby enable subsequent effective governance conversations (*4*, *5*).

However, witnessing both the fourth and sixth UNEA negotiations (UNEA-4 in 2019 and UNEA-6 in 2024) firsthand made it clear that the current “science-first, governance later” approach rests on a fundamental misunderstanding of the international politics of SRM. In the absence of effective or even minimal governance, accelerating unilateral and commercial research has accompanied widening divisions. By UNEA-6, the international divisions seen at UNEA-4 (*6*, *7*) between those who see SRM as potentially complementing more gradual decarbonization and countries that see it as a distraction from fair climate policy had only deepened. Moreover, the cast had expanded substantially. In 2019, the European Union (EU) (flanked by Bolivia) promoted precautionary governance. In 2024, many more developing nations came prepared to play a role, with African countries (as a coordinated bloc) leading the charge. States promoting more scientific research before governance looked more isolated, with the EU left to play the broker, outflanked by much of the Global South (see the figure).

Meanwhile, private-sector attention is growing with start-ups like Make Sunsets in the US and Stardust Solutions in Israel seeking to commercialize SRM. In the absence of international oversight, an escalating research race among leading states and corporations beckons, leaving the scientific, political, and ethical challenges that SRM poses, including its risk to sorely needed accelerated mitigation efforts, to fester. The prospects of a slippery slope to messy deployment seem greater than ever.

Deferring governance until considering deployment of geoengineering is proving an illusion that increases risks of any future deployment being poorly coordinated, counterproductive, and guided by narrow geopolitical or special interests. This would be worlds away from the optimized climate risk-management scenarios underpinning much geoengineering modeling, which leave aside critical questions about how to decide purpose and extent and how to implement, monitor, attribute, and control SRM (*8*). Without governance to achieve all this, optimistic model findings are not remotely realizable in practice. National governance of research programs (*9*) would, of course, be welcome. Yet to be effective, governance must extend beyond any single nation or bloc.

### THE DEADLOCK

Our accounts of the UNEA meetings—the most prominent intergovernmental discussions of solar geoengineering to date—are based on our direct observation of the negotiations, conversations with delegates and other observers, and detailed reading of the draft texts and submissions made by countries and accredited organizations as posted on the UN resolutions portal (*10*). Our observations at these meetings indicate that a scarcity of scientific and technical knowledge is not what prevents governance. Countries were deeply divided over Swiss-led draft resolutions, disagreeing on how or whether SRM development should be enabled or constrained, and over what counts as valid evidence, exposing diverging world views and interests, not least regarding the speed of emissions cuts.

The deadlock exists primarily between two groups of states distributed, unevenly, along a continuum (see the figure). The vast majority—around four-fifths of those participating or represented in the talks—hold or tend toward a “skeptical,” precautionary stance regarding SRM. They see it as riddled with uncertainty and fear that ungoverned pursuit of it will delay, disrupt, or undermine the accelerated international action urgently needed to cut greenhouse gas emissions. They want existing fragments of global governance of SRM to be acknowledged and a broader set of knowledges—beyond natural science and encompassing the governance risks and controversies that matter most to them—to be considered.

A much smaller but globally influential group of “permissive” states emphasize keeping SRM options open as a possible way to curtail rising temperatures, thus easing the urgency of decarbonization. At UNEA-4, the US and Saudi Arabia argued for “science first” and that the Intergovernmental Panel on Climate Change (IPCC), not the UN Environment Programme (UNEP), should lead any international assessment of SRM. They saw biophysical science about climatic implications—obtained through modeling or experimentation—as the overriding priority. They protested that even discussing governance internationally would discourage such research, and they demanded that SRM be detached from discussion of carbon dioxide removal

---

[1] Emmett Institute, School of Law, University of California Los Angeles, Los Angeles, CA, USA. [2] School of Politics and International Studies, University of Leeds, Leeds, UK. Email: duncan.p.mclaren@gmail.com

(CDR). By UNEA-6 in 2024, the IPCC had included SRM in its Sixth Assessment Report and the Swiss had excluded CDR from their new proposal. Still, the "permissives" argued once again for "science first" before any steps toward international governance. Apparently unsatisfied by the IPCC's (somewhat critical) assessment of SRM, this time they posited the new geoengineering "lighthouse" activity of the World Meteorological Organization's (WMO's) World Climate Research Program (WCRP) as the venue to compile more scientific knowledge, in preference to the more holistic and governance-oriented UNEP. They again resisted efforts to widen the types of knowledge with which SRM would be evaluated, prioritizing science about biophysical climatic effects over social and political scholarship. Knowledge generated by the humanities or from Indigenous ways of knowing was given short shrift.

Skeptical states sensed a slippery slope and rejected efforts at compromise involving aspects of the "permissive" positions they had already repeatedly rejected. Result: deadlock.

## CLASHING INTERESTS

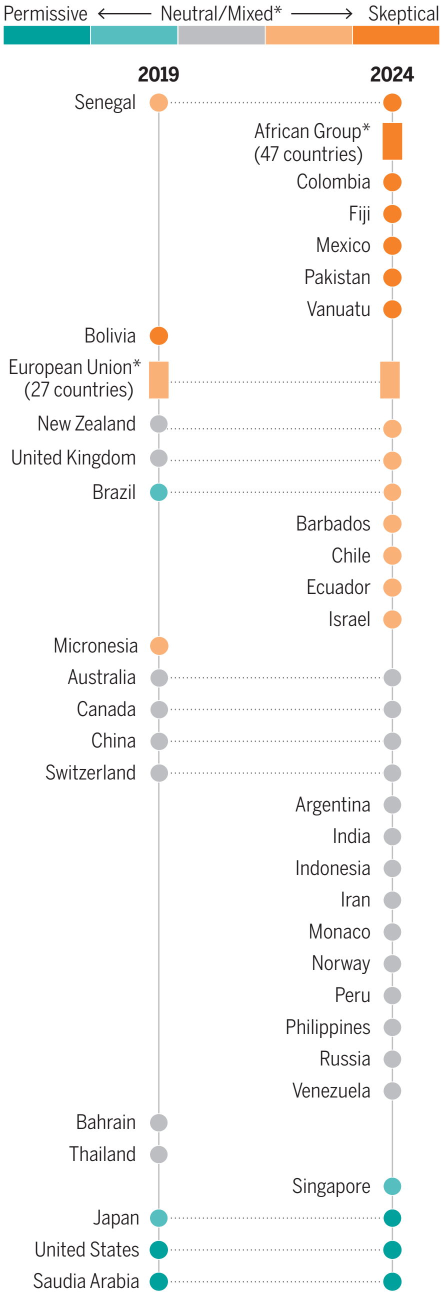
Rather than a paucity of science, such disagreements reflect diverging values and interests, and related negotiating stances. Or as one delegate pointedly put it, "every country has a right" to a worldview, not just the West. Country positions must be seen in the context of the underlying (geo)political economy of the global climate debate, and competing national and economic interests, whether in prolonging the fossil era or securing climate finance and Northern emissions cuts.

Although several accredited nongovernmental organizations were present at UNEA or submitted comments and position statements online and through the UN's "major groups" process, and most supported a skeptical position—some specifically citing a widely publicized, although contested academic proposal for a "non-use agreement" (*11*)—others actively advocated for accelerated scientific work and endorsed the WMO over UNEP. Hence, it would be facile to infer that country positions were determined by such lobbying.

But neither were positions determined by scientific estimates of future climate impacts and technical judgments about whether SRM alone could avert them. Mexico invoked "sovereignty," calling for transparency and "deploring" unauthorized outdoor experiments after balloon launches by the US-based Make Sunsets on Mexican territory. Whatever their motivations, the US and Saudi Arabia introduced language more accommodating of SRM and commensurate with an interest in slower or more "orderly" [in the language advocated at the 2023 UN climate conference (COP28)] decarbonization. Both major fossil producers and exporters, they pushed for reference to possible "benefits" as well as "risks" of SRM, rejecting more precautionary language favored by other states. They resisted or weakened proposed preambular text "recalling" existing intergovernmental references to geoengineering governance elsewhere, e.g., declarations of parties to the Convention on Biological Diversity and an African Ministerial Conference call for a "non-use" framework. They inserted wording about "temperature targets," which positions SRM more favorably as a "solution," or as a partial replacement for emissions reductions. They also proposed references to risks of abrupt climate change as a justification for considering SRM, even though the recent Global Tipping Points review rejects reliance on SRM as a tool to avoid tipping events (*12*).

### Growing skepticism

Positions ranging from "skeptical" to "permissive" were qualitatively allocated for all countries that submitted views or proposed text, based on the content, tone, and frequency of interventions.



*Neutral/Mixed: Countries expressing partial or diverging views on different aspects, and whose delegates attended negotiating sessions but made no interventions; European Union: All countries belonging to the European Union as of February 2024; African Group: African UN states, excluding those suspended from the African Union as of Feburary 2024. Senegal is the only country in Africa with a 2019 position and is represented separately.

GRAPHIC: M. HERSHER/*SCIENCE*

By contrast, many of the skeptical states explicitly rejected abstract "risk-benefit" approaches to assessment, refusing to afford equal weight to speculative "benefits." Favoring language of "risks and concerns," they applied precaution to SRM itself—rather than seeing it as a possible precautionary tool to tackle climate change—understanding that benefits of SRM seen in models are not reproducible in real-world applications if technical, political, and other conditions do not apply (*8*). They sought to insert wording referring to prior precautionary multilateral decisions and favored the UNEP as a forum. Strikingly, some of the countries that are most vulnerable to climate change—including Pakistan and Fiji—ultimately dismissed SRM as a credible climate response.

The US routinely exempts itself from international legal commitments and restrictive norms, leaving it a freer hand—in this case to pursue SRM research as it sees fit, perhaps to protect its preeminent global position and a perceived US technological lead [as one (non-Western) delegate suggested to us]. But by 2024, it was clear that UNEA is part of an increasingly multipolar international order in which middle-income and smaller states have more clout and confidence.

Compromise drafts were tabled that incorporated permissive and some skeptical points, although notably never including language on "non-use." Neither group was happy. Even when the text was pared back to merely "exploring options" for setting up a "repository" for voluntary collation of existing scientific information, states disagreed vociferously over what sort of research, and how and where findings would be gathered and validated. Such a process was seen as legitimizing more scientific research that some countries viewed as urgent and others as premature and risky, greasing an already slippery slope.

## GOVERNANCE FIRST?

Pushing on with unilateral research on the presumption that skeptical views are ill-formed opinions that will be changed simply by more science or information seems foolhardy (given the long and polarized history of climate controversies). A research-first approach would also be hard to justify as a moral duty to the most vulnerable countries as long as these adamantly oppose ungoverned SRM development.

> **“Given that premature deployment is almost universally ruled out, why not start there?”**

Yet philanthropic and public research funding is ramping up in the US and UK, and at small scale in some developing countries. Existing research governance mechanisms in academia are limited in scope. Voluntary measures so far have proven ineffective at seeding ethical global norms around SRM research and development, defusing competition, or reining in commercialization. Proposals involving consistent national-level research governance [e.g., (*9*, *13*)] remain contested among scientists, and developing national governance unilaterally to enable more research focused on the biophysical climate implications would risk elevating the sort of controversy seen at UNEA.

However, what seems to have been ignored is that the vast majority of participants at UNEA indicated they would view moves to deploy as wholly illegitimate. SRM scientists are typically adamant they are proponents of research—not deployment. Commercial actors claiming to have performed SRM operations are roundly denounced as “rogue” and “irresponsible” (*14*). Both the international “Climate Overshoot Commission” and the European Commission’s Chief Scientific Advisors (*15*) have called for a moratorium on deployment and large-scale experimentation.

Given that premature deployment is almost universally ruled out, why not start there? This common ground could be seized on by states and formalized as an international nondeployment framework. This would immediately discourage rogue ventures from uncoordinated deployments and temper corporate profit-seeking, slowing any competitive research “arms races.” Rather than acting as a de facto research ban but implemented as a scaffold within which principles for responsible and ethical SRM research might be negotiated, such a framework could assuage some slippery-slope concerns while potentially enabling well-governed research. We use the term “nondeployment” as distinct from “non-use” as framed in the “non-use agreement” proposal (*11*), which also seeks to preclude outdoor experimentation, public funding for development, international assessment, and patenting of SRM. In practice, nondeployment might take the form of an indefinite moratorium and could cover diverse SRM techniques with global or transboundary impact, subject to negotiations. Principles for responsible research, such as transparency, shared intellectual property, integration of justice concerns, independent ethical review of proposals, and advance public and stakeholder engagement [e.g., (*13*)], might be codified. Further SRM modeling—prioritized by certain states—could be reoriented toward nonideal scenarios and situated within the broader scrutiny of environmental, legal, geopolitical, and security implications requested by other states. More widely, research about how to negotiate a world where Paris Agreement goals are not met might progress beyond “overshoot” techno-solutionism to engage the broader and deeper global concerns that we heard raised at UNEA.

Entrenched positions on research governance, and the international dynamics behind them, may obstruct a nondeployment framework too, of course. But skeptical countries should welcome it, and the “costs” for permissive countries of an agreement prohibiting deployment should be very low (or negative)—unless they actually are ready to see deployment, despite protesting that it is premature to even discuss governance. This seems unlikely but would naturally change (or clarify) the terms of debate considerably.

Some past proposals for governance of SRM research have been interpreted by critics as seeking to prevent science and framed as “willfully neglecting” SRM (*4*) or “choosing not to know” (*3*). But this would mischaracterize skeptical states at UNEA. They did want to know more about geopolitical, ethical, legal, social, and wider environmental dimensions but were concerned that more natural science in isolation—and prior to governance—would exacerbate the types of risks and uncertainties that SRM raises, which are not only biophysical but also social and geopolitical (*2*). The EU and many developing nations repeatedly demanded that any assessment should prioritize understanding all the possible risks of SRM—and require a broad knowledge base. A narrow approach that frames SRM as desirable as long as “the science” deems it so raised concerns. Skeptical states (especially Latin American ones) also wanted to see more equitable access to existing knowledge, currently dominated by a small group of Northern nations, before further scientific experimentation.

In theory, then, well-governed broad-spectrum research, embracing diverse knowledge types, should help everyone better understand both SRM and climate issues in general. Yet at both UNEA meetings, permissive states argued a resolution including such provisions would “preempt the science”—unilateral science that others saw as already unbalanced and risky. In this context, governance of deployment provides a—perhaps fleeting—route to arresting the current slide toward uncoordinated development and commercialization. But the further permissive countries push down a slippery slope, the more skeptical opposition accrues, the less chance of a formal agreement on nondeployment. And without that, coordinated and well-governed research efforts directing holistic and responsible inquiry for collective benefit seem dead in the—steadily acidifying—water. ■

### REFERENCES AND NOTES

1. Royal Society, *Geoengineering the Climate: Science, Governance and Uncertainty* (Royal Society, 2009).
2. T. Felgenhauer *et al*., *Solar Radiation Modification: A Risk-Risk Analysis* (Carnegie Climate Governance Initiative, 2022).
3. E. A. Parson, H. J. Buck, S. Jinnah, J. Moreno-Cruz, S. Nicholson, *Wiley Interdiscip. Rev. Clim. Change* **15**, e903 (2024).
4. C. E. Wieners *et al*., *Oxford Open Clim. Change* **3**, kgad002 (2023).
5. J. E. Aldy *et al*., *Science* **374**, 815 (2021).
6. S. Jinnah, S. Nicholson, *Nat. Geosci.* **12**, 876 (2019).
7. D. McLaren, O. Corry, *Glob. Policy* **12** (S1), 20 (2021).
8. D. McLaren, *Energy Res. Soc. Sci*. **44**, 209 (2018).
9. S. Jinnah *et al*., *Science* **385**, 600 (2024).
10. These detailed materials are made available to accredited participants but may not be circulated or published elsewhere. We hold a set of these papers downloaded from the UN portal, which were made available as part of the review process and may be viewed on request, or accessed through independent accreditation. A summary of the negotiations reported by the Third World Network may also be consulted online at https://www.twn.my/title2/climate/info.service/2024/cc240302.htm (last accessed 1 November 2024).
11. F. Biermann *et al*., *Wiley Interdiscip. Rev. Clim. Change* **13**, 754e (2022).
12. T. M. Lenton *et al*., Eds., “The Global Tipping Points Report 2023” (2023); https://global-tipping-points.org/resources/.
13. American Geophysical Union, “Ethical Framework Principles for Climate Intervention Research” (2024); https://doi.org/10.22541/essoar.172917365.53105072/v1.
14. M. S. Diamond, K. Wanser, O. Boucher, *Clim. Change* **176**, 96 (2023).
15. European Commission, Directorate-General for Research and Innovation and Group of Chief Scientific Advisors, “Solar radiation modification,” Publications Office of the European Union (2024); https://data.europa.eu/doi/10.2777/391614.

**ACKNOWLEDGMENTS**

The authors thank the Friends World Committee for Consultation for enabling our participation in UNEA-6 as part of their delegation.

10.1126/science.adr9237

BOOKS *et al.*

PSYCHOLOGY

# Data crumbs and sitting ducks

## Our digital footprints offer potentially lucrative personality insights to the highest bidders

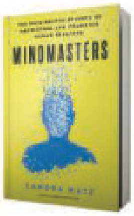
By **Jonathan Wai**

Sandra Matz grew up in the town of Vögisheim, population 500. The 499 other residents knew numerous details about her life and used this information to infer her thought processes and to offer unsolicited advice. Information, she learned, can be a form of control.

In *Mindmasters*, Matz—now a computational social scientist—uses this analogy to explain that nearly all of us are offering others insights into our personality, interests, abilities—our human individuality (*1*)—through the data crumbs we leave behind, all of which can be sold to the highest bidder (who usually wants to sell us something). The problem, Matz argues, is that we do not seem to realize that we are being targeted. We passively agree to share our data all the time without much resistance, and with advances in artificial intelligence (AI), third-party surveillance will likely get worse (*2*). The core arguments in this book are not new (*3*), but Matz writes very clearly and makes a compelling and current case for taking seriously the threats posed by data-driven psychological targeting.

The reviewer is at the Department of Education Reform and Department of Psychology, University of Arkansas, Fayetteville, AR, USA. Email: jwai@uark.edu

Part one of *Mindmasters* looks at data as a window into our psychology, specifically within the context of Matz's research around the "Big Five" personality traits: openness, conscientiousness, extroversion, agreeableness, and neuroticism. Each of us generates six gigabytes of data every hour, and Matz explains that our Facebook likes, geolocation data, Google searches, and even our pictures can be used to make inferences about our personalities that can then be used for targeted marketing. (Credit card transactions are supposed to be anonymous, but it turns out that knowledge of just three purchases can enable an actor to infer a purchaser's personal identity.)



**Mindmasters: The Data-Driven Science of Predicting and Changing Human Behavior**
*Sandra Matz*
Harvard Business Review Press, 2025. 240 pp.

In the book's second part, Matz examines the positive and negative sides of psychological targeting. Here, she makes it clear that although "psychological targeting won't easily change people's core identities," this does not mean that people cannot be nudged in meaningful ways when making important decisions. Moral reframing in political advertisements, for example, "can sell the same political idea to different people."

Matz thinks the potential positives of targeted marketing might include better access to resources that could help us improve our health, engage politically, learn more effectively, or more easily find a job we will love. But our data could also be used by suppressive regimes to predict whether an individual will join a protest or petition for government action or as a means of health surveillance that could restrict the movements of those with certain illnesses or lead to discrimination against those with disabilities.

We make roughly 35,000 decisions daily, Matz notes, and "we are influenced by what other people do, the context in which a decision is made, or our mental state at the time we make a decision." However, her research suggests that despite our susceptibility to influence, people just do not seem all that concerned about data privacy. But perhaps this is not surprising. Most of us simply do not have the capacity, time, and energy to manage our data privacy.

Matz argues forcefully that we need a data ecosystem that prioritizes our personal security and makes limiting third-party access to our data easy and, ideally, automatic. A default "opt-in" system for organ donation makes a remarkable difference, and it is reasonable to assume that a default "opt-out" for data privacy would do much the same. The problem, as Matz and others (*4*) note, is that such policies are not in the interest of the big tech companies and others who feed on the existing ecosystem.

"No matter how safe and comfortable you feel now, your data could be misused in the future," writes Matz. Those interested in a more extensive consideration of the potential future perils posed by data generated today should consult Carl Öhman's *The Afterlife of Data* [(*5*); and see the *Science* review (*6*)].

The moral and ethical questions surrounding data are important. In *Mindmasters*, Matz focuses on her own work on inferring personality traits from data, but the broader literature suggests that data-driven marketing will likely be able to infer (and thus use) many more aspects of who we are to target us. Looking at history can help us understand ways we might balance innovation with the public good (*2*). Instead of acquiescing to data collection schemes mediated by byzantine user agreements and defaults that prioritize the interests of tech giants, we need to take back control of our data, especially with the rise of AI and the potential for its misuse in this context (*7*). ■

### REFERENCES AND NOTES

1. W. Revelle, J. Wilt, D. M. Condon, in *The Wiley-Blackwell Handbook of Individual Differences*, T. Chamorro-Premuzic, S. von Stumm, A. Furnham, Eds. (Blackwell Publishing, 2011), chap. 1.
2. T. Wheeler, *Techlash* (Brookings, 2023).
3. C. Rudder, *Dataclysm* (Crown, 2015).
4. C. Véliz, *Privacy Is Power* (Bantam Press, 2021).
5. C. Öhman, *The Afterlife of Data* (Univ. of Chicago Press, 2024).
6. K. Shilton, *Science* **384**, 392 (2024).
7. G. Marcus, *Taming Silicon Valley* (The MIT Press, 2024).

10.1126/science.adt5945

ILLUSTRATION: DAVIDE BONAZZI/ISPOT STOCK

AUTOMATION

# Hidden human costs of AI

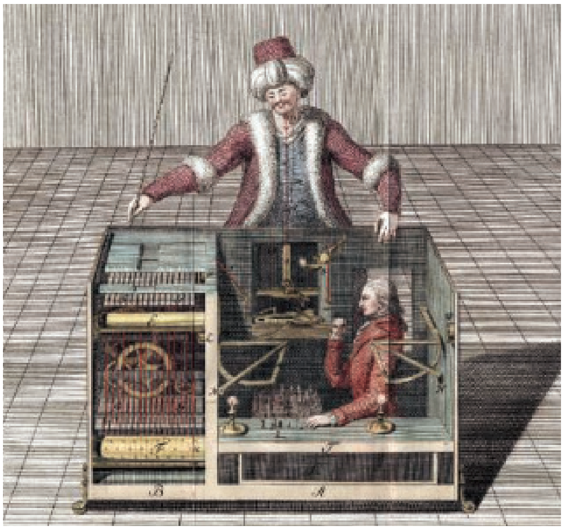
## A sociologist interrogates the invisible labor that underlies "autonomous" systems

*By* **Dov Greenbaum**[1,2,3,4] *and* **Mark Gerstein**[4,5,6]

**Waiting for Robots:**
**The Hired Hands of**
**Automation**
*Antonio A. Casilli*
University of Chicago Press,
2025. 336 pp.

Sociologist Antonio Casilli's newly updated 2019 book *Waiting for Robots*, translated into English by Saskia Brown, offers a nuanced critique of automation and artificial intelligence (AI). In it, Casilli leverages his academic expertise to expose the intricate, often invisible human labor that supports these systems. At its core, the book challenges the idea of fully autonomous AI.

Structured into three parts, the book systematically unpacks the complexities of digital labor. The first section critiques the idea that AI eliminates human work, suggesting instead that it reconfigures labor in ways that obscure its value. The second section provides vivid examples of how digital platforms exploit three kinds of human labor: "on-demand work, microwork, and social media work." The final section examines the global inequalities perpetuated by these systems, proposing solutions such as platform cooperatives and redistributive income models.



Hidden human effort has a long history in automation.

Central to Casilli's argument is the concept of "artificial artificial intelligence," a term he uses to describe how "human input is a major component of all artificial intelligence tools." Platforms such as Amazon Mechanical Turk, Uber, and Facebook, for example, extract value from human effort through tasks such as data annotation, content moderation, and algorithm training. Casilli links this "microwork"—fragmented, small-scale, and often low-paying forms of labor that obscure the broader purpose of tasks while perpetuating the illusion of fully autonomous technology—to historical patterns of geopolitical exploitation.

Microwork, he writes, reveals a "new international division of labor"—hidden, underpaid, and unstable—where jobs in the Global North are not necessarily being automated by machines but are instead outsourced to digital migrants in the Global South. By connecting the precarious conditions of hidden workers to these systemic patterns, Casilli sheds light on the vulnerabilities and inequities underpinning the digital economy.

Building on this idea, the book next examines multisided markets, where platforms mediate interactions to extract value. Casilli offers insights into how video platforms monetize viewers by selling advertising space, often without compensating content creators directly. Similarly, he dissects Uber's surge pricing model, which raises fares during peak demand periods. While this system incentivizes drivers, it also provides Uber with valuable behavioral data, illustrating the asymmetry of value distribution in digital economies, where platforms benefit disproportionately from user participation and labor. (For scientists, Casilli's critique extends naturally to the scientific publishing ecosystem, where peer reviewers contribute unpaid labor that is essential to maintaining the rigor of scholarly research.)

The book explores how platforms blur the line between work and play, a phenomenon referred to as "playbor." Gamified tasks such as reCAPTCHA exemplify this trend, as participants unknowingly engage in unpaid labor disguised as simple, interactive activities. While Casilli questions the appropriateness of the term, observing that it "disregard[s] the fact that platform work can be tedious and time-consuming," his analysis raises critical ethical questions about manipulation and the erosion of informed consent in the digital age.

Beyond labor, Casilli examines the commodification of user data and metadata, uncovering how platforms extract value from records of consumer behaviors, interactions, and social connections. This practice raises considerable concerns about surveillance, privacy, and the exacerbation of global inequalities.

Although Casilli effectively critiques the heavy reliance on human labor in AI systems, he does not fully engage with recent technological developments in unsupervised and semisupervised learning. Innovations such as self-play learning, as exemplified by systems such as Google's AlphaZero, and the use of physically based data augmentation challenge the assumption that AI automation remains fundamentally dependent on manual labeling. Acknowledging these advancements could have added nuance to Casilli's argument, providing a more comprehensive view of the evolving interplay between human labor and machine learning.

The book concludes with a thought-provoking proposal for collectivist data ownership that aims to address the imbalances and exploitation inherent in platform capitalism's approach to data. Here, Casilli suggests that "data could become the direct, indivisible, and inalienable collective property of its users" and encourages readers to envision alternative structures for a fairer and more inclusive digital economy.

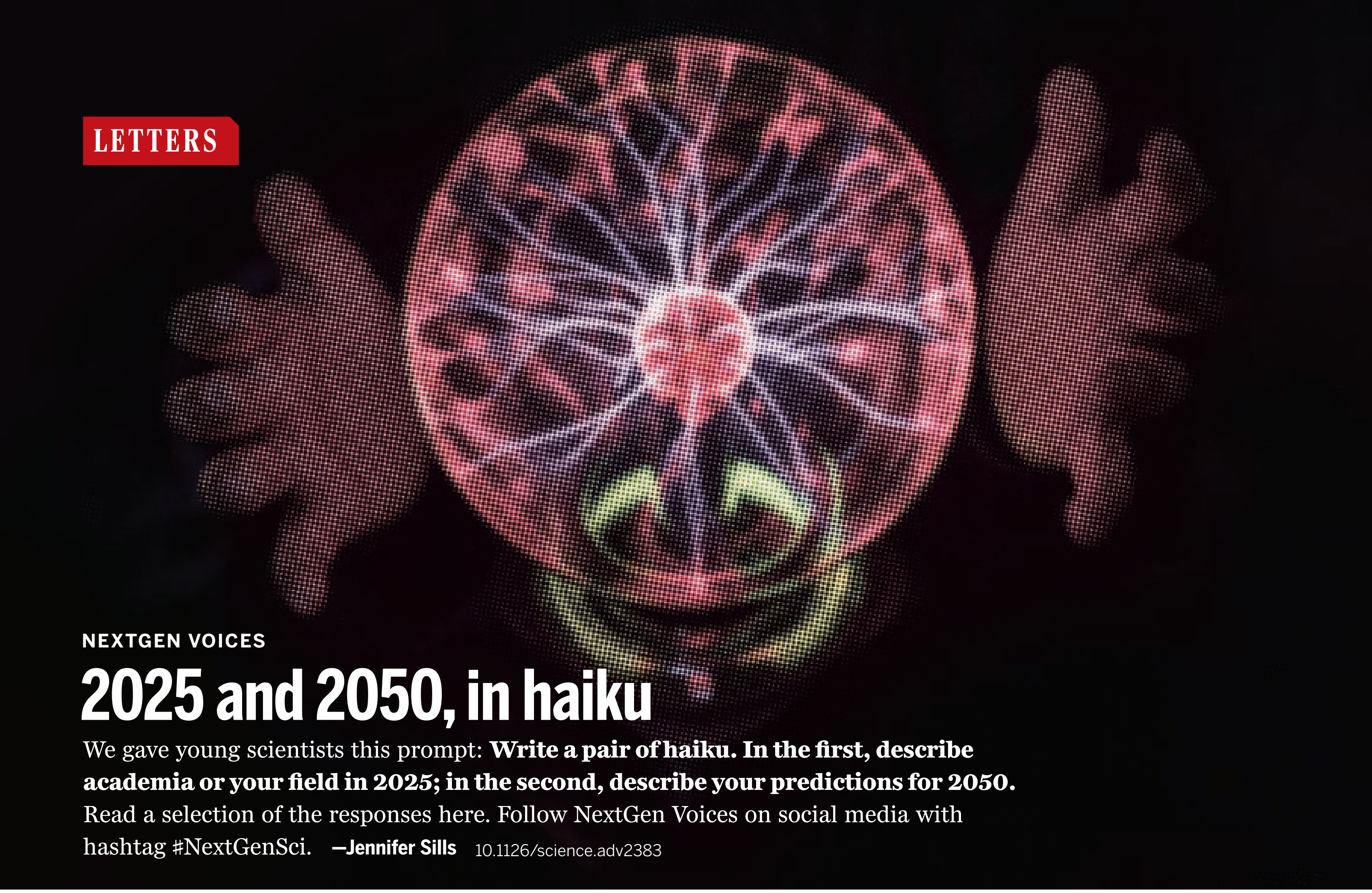
Overall, *Waiting for Robots* makes a substantial contribution to the discourse on digital labor and platform capitalism. Casilli's incisive analysis of the hidden human costs of AI, combined with his clear and engaging prose, establishes the book as essential reading for those seeking to understand the ethical and social dimensions of technology. By dismantling the conventional wisdom surrounding AI and exposing the labor that underpins it, Casilli offers a compelling framework for envisioning a more equitable and inclusive digital future. ■

10.1126/science.adu1541

ILLUSTRATION: JOSEPH RACKNITZ

---

The reviewers are at the [1]Zvi Meitar Institute for Legal Implications of Emerging Technologies, [2]Harry Radzyner Law School, and [3]Dina Recanati School of Medicine, Reichman University, Herzliya, Israel; and the [4]Department of Biomedical Informatics and Data Science, [5]Program in Computational Biology and Bioinformatics, and [6]Department of Computer Science, Yale University, New Haven, CT, USA. Email: dov.greenbaum@aya.yale.edu



LETTERS

NEXTGEN VOICES

# 2025 and 2050, in haiku

We gave young scientists this prompt: **Write a pair of haiku. In the first, describe academia or your field in 2025; in the second, describe your predictions for 2050.** Read a selection of the responses here. Follow NextGen Voices on social media with hashtag #NextGenSci. **—Jennifer Sills** 10.1126/science.adv2383

## Science and society

| 2025 | 2050 |
|---|---|
| **Review checklist**<br>By White, right? Accept!<br>No: This is by Miss Blanco<br>"Improve the writing." | **Review checklist reviewed**<br>By Blanco? Accept!<br>No: This is AI-written<br>"Improve the science."<br>**Rayane da Cruz Albino,** Brazil |
| **Open for business**<br>Data shared with all<br>Transparency pays off big<br>Who's cashing in most? | **Open secrets**<br>Research "free," but wait:<br>Corporations own the vaults<br>Access comes with fees<br>**Yana Suchikova,** Ukraine |
| **What is true?**<br>Fake news spreads so fast<br>Politicians pull the strings<br>Truth fights to break free | **Shared truth**<br>Barriers come down<br>Data shared without a price<br>Info shared with all<br>**Gustavo R. Gameiro,** Brazil |
| **Caffeinated**<br>Manuscripts pile high,<br>Coffee spills, but we still try<br>Mop it up, focus | **Automated**<br>AI grows, expands<br>Quantifying all we do<br>Where's the "Wheee!" factor?<br>**Shicong Feng,** China |

| 2025 | 2050 |
|---|---|
| **Copywrong**<br>All human knowledge<br>Without consent or credit<br>Used to train AI | **Copyright 2.0**<br>A call for justice<br>AI accuses humans<br>Of plagiarism<br>**Sarah Masud,** Denmark |
| **Obsolescence**<br>AI runs smart grids<br>LLMs are everywhere<br>Replacing humans? | **Nostalgia**<br>Singularity<br>Achieved—humans optimized<br>What of our true spark?<br>**Wadim Strielkowski,** Czech Republic |
| **Impressions**<br>Our economy<br>Carves unintended patterns<br>In the Earth's surface | **Decisions**<br>Will we ask for more<br>Now that we have what we want?<br>Or give back at last?<br>**Markus Vihma,** Estonia |
| **Unbelievable**<br>Hadrons collided<br>Then muons and dark matter<br>Sweet understanding | **Unbelieved**<br>On the cusp of all<br>A unified universe<br>But all trust is gone<br>**Bryan J. Field,** USA |

IMAGE: RAWPIXEL/ISTOCK PHOTO. ADAPTED BY N. JESSUP/SCIENCE

# Human health

### 2025 / 2050

**Planting hope**
World of AI grows
Hopeful beginnings take root
Secrets to health bloom

**Programming hope**
AI maps our genes
Cures from technology rise
Flesh and code entwined

**Rida Ismail,** Cyprus

**Who is cured?**
More cures discovered
Who benefits from progress?
Time and place decide

**Ethics obscured**
Genes reimagined
Humans altered as deemed best
GPT decides

**Aishwarya Mitra,** USA

**Owning drugs**
New medicines made
Who will own the key to health?
Fight over patents

**Owning destinies**
Can we change their fates?
Change lives before they are born?
Is this advancement?

**Julia Gauffreau,** USA

**Family blueprint**
Ancestral whispers
Map cancer's unique defense
Precision's new guide

**Personal plan**
No more one-size cure
Genes craft each unique pathway
Healing, tailor-made

**Howard Lopes Ribeiro Junior,** Brazil

### 2025 / 2050

**Organoids**
Cells grow into form
Organoids spark life anew
Bridging hope and cures

**Organs on demand**
Tiny cells take shape
Livers, hearts, grown while you wait
Cure or clone? Debate!

**Shihao He,** China

**Search for knowledge**
Plaques pile up in brains
Mysterious, haunting clumps
Do they hold the key?

**Knowledge overload**
Machines know too soon
The mind's fate sealed in code lines
Dread before decline

**Junle Chen,** USA

**Teaching evolution**
Famous Scopes Trial
100 years in the past
Have we evolved yet?

**Evolution evolves**
CRISPR everywhere
My children, my food, my cat
"All natural," all!

**Prosanta Chakrabarty,** USA

**Antibiotic setback**
Microbes learn and change
Antibiotics falter
Lives lost everywhere

**Phages forge ahead**
Hunters of microbes
Phages conquer where drugs fail
New hope, lives reclaimed

**Song Lin Chua,** Hong Kong

# Sustainable strategies

### 2025 / 2050

**Power grows**
Batteries grow strong
Cars run far, charge in minutes
Future feels so near

**Flowers grow**
Recycling makes sense
Batteries live many lives
Earth's health benefits

**Zhen Wang,** China

**Green transition**
Winds of change arise
Solar beams ignite the grid
Carbon shadows fade

**Net-zero success**
Smart currents align
Sun and wind in perfect dance
Resilience achieved

**Zelong Lu,** China

**Suspense**
Turning waste to fuel
Will we choose green or profit?
Earth waits, holds its breath

**Relief**
Solar cells bloom bright
Water splits, gas turns to fuel
Tomorrow shines green

**Shuoren Li,** China

**Carbon cycle**
Building carbon sinks
And addressing climate change
Already too late?

**Carbon conversion**
Processed exhaust gas
Becomes fuel, clothing, and food
Some more carbon, please!

**Yuan Zhi,** China

# Smart technology

### 2025 / 2050

**Tiny solutions**
Small bits heal our cells
Yet where do nanodusts go?
Unseen risks linger

**Nano-profusions**
Tiny minds compute
Atoms learn and self-repair
Machines small as thought

**Karen Jacqueline Cloete,** South Africa

**Stronger fields**
CRISPR cuts with care
Plants fight unseen viral wars
Fields wait, genome-strong

**Better yields**
AI scouts the field
Sensors track each root's heartbeat
Life, data-driven

**Sara Shakir,** France

**Skilled plants**
Plants, nature's filter
Sipping toxins, growing strong
A verdant hope shines

**Synthetic leaves**
Solar cells in leaves
A new photosynthesis
Nature, tech entwined

**Ankita Gupta,** India

**Silicon lemon**
Squeezing the last drop
Of that silicon lemon
Howling, "Moore, please more!"

**Chip lemonade**
2D layers bloom
Silicon's epoch withers
Chips thrum, "Moore, no more!"

**Ufuk Erkılıç,** Japan

# RESEARCH

## IN *SCIENCE* JOURNALS

*Edited by* **Michael Funk**

**GLOBAL WARMING**

### Reflecting on climate change

Why was the year 2023 so much warmer than expected? Anthropogenic forcing and El Niño have been suggested as at least part of the reason, but they cannot account for the magnitude of the temperature jump. Goessling *et al*. identify another cause: a record-low planetary albedo caused mainly by reduced low-cloud cover in the northern mid-latitudes and tropics. If this shift represents an excursion into a new normal, our future could be hotter faster than expected. —Jesse Smith *Science* p. 69, 10.1126/science.adq7280

Sunlight reflects off clouds and aerosols over the Pacific Ocean.

PHOTO: IMAGE SCIENCE AND ANALYSIS LABORATORY, NASA-JOHNSON SPACE CENTER.

**IMMUNE REGULATION**

### The dark side of immunotherapy

Immune-related adverse events (irAEs) are side effects of immune checkpoint blockade (ICB). Borys *et al*. investigated how ICB can cause immune dysregulation in the submandibular gland and contribute to irAEs in mice. Deficiency of the programmed cell death 1 (PD-1) receptor or treatment with an antibody against its ligand, PD-L1, resulted in the expansion of CD8$^+$ T cells in the submandibular gland. PD-L1–expressing natural killer cells were required to regulate a population of CD8$^+$ T cells that phenotypically resembled the pathogenic GzmK$^+$ CD8$^+$ T cells seen in Sjögren's syndrome patients. These findings demonstrate that natural killer cells play a critical role in controlling CD8$^+$ T cell expansion in the submandibular gland. —Hannah M. Isles

*Sci. Immunol.* (2024) 10.1126/sciimmunol.adl2967

**ELECTRONIC MATERIALS**

### Improving ultrathin conductors

Amorphous semimetal niobium phosphide has greater surface conductance as nanometer-scale films than the bulk material and could enable applications in nanoscale electronics. Khan *et al*. grew amorphous thin films of niobium phosphide—a material that is a topological insulator as a crystalline material—as nanocrystals in an amorphous matrix. For films with 1.5-nanometer thickness, this material was more than twice as conductive as copper. —Phil Szuromi

*Science* p. 62, 10.1126/science.adq7096

**EVOLUTIONARY BIOLOGY**

### Luck of the draw

The world, both natural and human, is full of inequities, and researchers have long desired to understand to what degree differences arise from "luck" (otherwise called contingency). Testing for this effect is extremely challenging across all systems given the variability of environments an organism is born into. Zipple *et al*. performed an experiment that equalized both the environment and genotype in mice within a semiwild, yet consistent, enclosure. They found that small contingent events were amplified across a mouse's lifetime, leading to competitive advantage overall. However, this effect only occurred in male mice, who compete for resources, whereas it was not seen in females, who do not. —Sacha Vignieri

*Science* p. 81, 10.1126/science.adq0579

**ORGANIC CHEMISTRY**

### Gently sliding in nitrogen

There has been a recent surge in reported methods for editing molecular frameworks one or two atoms at a time, in the interest of diversifying pools of compounds synthesized for drug discovery campaigns. Ghosh *et al*. now report a comparatively convenient and versatile method to insert nitrogen into five-membered pyrrole, indole, and imidazole rings. The reaction hinges on sulfenylnitrenes derived from stable precursors. The sulfur-based leaving group obviates the need for oxidants and so renders the process compatible with oxidation-sensitive functionality. —Jake S. Yeston

*Science* p. 102, 10.1126/science.adp0974

PHOTORECEPTORS

## Ultraviolet light protection in fish

Photoreceptors have been recently shown to be expressed in nonvisual organs. Their role remains to be fully elucidated. Fukuda *et al*. found that pituitary endocrine cells of Japanese rice fish respond to light stimulation. Exposure of melanotrophs to ultraviolet light increased calcium influx in these cells, resulting in melanocyte-secreting hormone release. Fish lacking the photoreceptor protein Opn5m showed reduced melanogenesis, which suggests that nonvisual photoreceptors might serve as an ultraviolet light protection mechanism. —Mattia Maroso

*Science* p. 43, 10.1126/science.adj9687

SYNTHETIC BIOLOGY

## Designing phosphorylation-based circuits

Modification of proteins by covalent phosphorylation is a primary mechanism for signal processing in biological systems. Yang *et al*. developed a system to use the modularity of such systems for the design of synthetic biochemical signaling based on phosphorylation rather than genetic circuitry. Combining various protein domains, they designed synthetic kinase-, substrate-, and phosphorylation-dependent interaction domains to create synthetic circuits that could sense and amplify an external signal and activate expression of a transgene. For example, in human T cells, such a circuit could sense tumor necrosis factor–α and maintain controlled cytokine secretion. Rapid regulation and reversibility of phosphorylation-based signals and programmed protein-protein interactions offer the opportunity for versatile design and modeling of circuits that could be useful for a range of therapeutic and diagnostic applications. —L. Bryan Ray

*Science* p. 74, 10.1126/science.adm8485

ELECTROCHEMISTRY

## Stabilizing ruthenium oxide

The addition of tantalum to ruthenium oxide slows its dissolution during the oxygen evolution reaction for water splitting and can enable it to be a replacement for iridium oxide. Zhang *et al*. show that dissolution was sensitive to surface structure and that addition of tantalum slowed corrosion and increased activity. After 2800 hours of operation at a current density of 1 ampere per square centimeter, the catalyst $Ta_{0.1}Ru_{0.9}O_{2-x}$ exhibited almost constant voltage, with a decay rate less than ~14 microvolts per hour. —Phil Szuromi

*Science* p. 48, 10.1126/science.ado9938

COMPUTER SCIENCE

## Tuning particle accelerators with AI

Large language models (LLMs) can help scientists tune the sophisticated particle accelerator machines used in high-energy physics, cancer therapy, and material sciences. Kaiser *et al.* tested whether LLMs can autonomously tune accelerator subsystems using only written prompts from the machine's operator rather than the standard technical algorithms. In their tests, about half of the LLM systems that were evaluated showed promise at this task, even as existing optimization algorithms remain superior. This proof-of-concept study illustrates the potential for modern artificial intelligence (AI) to help scientists more rapidly solve complex numerical optimization tasks, which are increasingly common in modern science. —Aaron Clauset

*Sci. Adv.* (2025)
10.1126/sciadv.adr4173

# IN OTHER JOURNALS

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

DEEP-SEA LIFE

## A bizarre sea slug in the deep

Deep-sea biodiversity and the ecological adaptations of life are less explored in the bathypelagic zone than in other oceanic zones. Robison and Haddock describe *Bathydevius caudactylus*, a new nudibranch (sea slug) species that is found at depths of 1013 to 4009 meters in the eastern North Pacific Ocean. This species exhibits bioluminescence and distinctive anatomical features, including a gelatinous body with a bell-shaped oral hood and finger-like projections on its tail, which distinguish it from known species. This study highlights the importance of exploring deep-sea environments because they harbor previously unknown organisms that could provide insights into evolutionary biology and ecosystem dynamics. —Di Jiang

*Deep Sea Res. I Oceanogr. Res. Pap.* (2024)
10.1016/j.dsr.2024.104414

PHOTO: NOAA/MONTEREY BAY AQUARIUM RESEARCH INSTITUTE

INTRODUCED SPECIES

## Cat's out of the bag

Humans have spread the domestic cat (*Felis catus*) from North Africa to most of the world. Free-ranging cats kill many wild birds and mammals, especially on islands with naïve prey. Philippe-Lesaffre *et al.* used data on cat prey and machine learning to compare the diets of cats in Australia, where cats weren't introduced until 1788, with those on continents where they were introduced in the 16th century (Europe and North America). Small passerine songbirds with larger geographic ranges were common prey across continents. In Australia, parrots and many mammals were vulnerable to predation, and unlike on other continents, larger body size did not protect mammals from predation. This work provides further evidence that fauna are vulnerable to new predators. —Bianca Lopez

*Ecography* (2024)
10.1111/ecog.07169

NEUROIMMUNOLOGY

## No pain, no gain?

Signaling between macrophages and sensory neurons may guard against damaging molecules that enter joints during systemic inflammation. Hasegawa *et al.* identified a specific type of capillary within the leg joints of mice that allowed small molecules from the blood to enter the synovial tissue. Three distinctive macrophage subsets lined these capillaries and

A new species of sea slug, *Bathydevius caudactylus*, has been discovered in the Pacific Ocean.

were localized close to nociceptors, neurons that mediate pain. Each macrophage subset reacted to immune stimuli in the blood but had different responses. One type activated the nociceptors, partly through the production of the cytokine interleukin-1β. The nociceptors could in turn release calcitonin gene-related peptide (CGRP), a signal that promoted the other macrophage types to cluster and form a protective structure around the capillaries. —Sarah H. Ross

*Nat. Immunol.* (2024) 10.1038/s41590-024-02011-8

QUANTUM SIMULATION

## Painting the potential

Ultracold atoms are a promising platform for simulating quantum many-body phenomena that are difficult to access using classical computers. With this goal in mind, much progress has been made in making the range, shape, and sign of atom-atom interactions more tunable. These strategies have included using atoms and molecules with large dipole moments as well as placing atoms in optical cavities, with photons mediating interactions. Bonifacio *et al.* propose a cavity-based scheme that may take this flexibility a step further. In their proposal, a laser periodically scans an atomic cloud residing in a cavity; if the various relevant experimental timescales are favorable, this creates a fully tunable interaction potential. The technique may lead to the realization of exotic collective phases. —Jelena Stajic

*Phys. Rev. X* (2024) 10.1103/PRXQuantum.5.040332

GREEN ENERGY

## The greening of Texas

The state of Texas is the largest oil and natural gas producer in the United States, and the fuels it produces are the source of one-third of the total carbon emissions of the entire country. The decarbonization of this industry is a tall order, but also one with tremendous potential benefits to the environment and the economy. Lin *et al.* describe how, by using its extensive energy infrastructure and investment resources, Texas could become a major green hydrogen supplier, helping to decarbonize not only the United States but the rest of the world. —Jesse Smith

*Proc. Natl. Acad. Sci. U.S.A.* (2024) 10.1073/pnas.2321347121

BIOADHESIVES

## How ticks stick

Insect parasites have invented ingenious ways of attaching to and manipulating their hosts through salivary secretions. In addition to grasping with their mouthparts, ticks exude a proteinaceous cement that helps them stay attached over several days. Ganar *et al.* used in vitro peptide synthesis to study a glycine-rich protein from the deer tick *Ixodes scapularis*. They show that this protein forms liquid condensates and transitions to a gel-like adhesive over time. This mechanism may be common to other glycine-rich peptides in tick saliva or in other bioadhesives produced by invertebrates. —Michael A. Funk

*Nat. Chem.* (2024) 10.1038/s41557-024-01686-8

GUT MICROBIOTA

## Eat your polyphenols

A diet high in plant polyphenols is considered beneficial because of the anti-inflammatory effects of these molecules. Found in olive oil, coffee, various herbs, and fruit, hydrocaffeic acid is one of the most abundant polyphenols in the gut. However, its beneficial effects are reduced when it undergoes dehydroxylation. In a metatranscriptomic survey of human gut microbiome genes, Bae *et al.* discovered a previously unknown bacterial catechol dehydroxylase. The main source of the enzyme was traced to *Gordonibacter pamelaeae*, a relative of *Eggerthella lenta* that is well known for its drug-transforming properties. A reduction in genes encoding the enzyme correlated with lower levels of gut inflammation in people who ate green vegetables. —Caroline Ash

*Cell Host Microbe* (2024) 10.1016/j.chom.2024.10.002

REVIEW SUMMARY

CONSERVATION

# Loss of Earth's old, wise, and large animals

**R. Keller Kopf*, Sam Banks, Lauren J. N. Brent, Paul Humphries, Chris J. Jolly, Phyllis C. Lee, Osmar J. Luiz, Dale Nimmo, Kirk O. Winemiller**

**BACKGROUND:** Humans have caused a decline in old age classes of wild animal populations, whereby many of Earth's oldest, often largest, and most experienced individuals have been eliminated from ecosystems. The ecological and socioeconomic consequences of this loss are poorly understood. Owing to the well-established negative aspects of biological aging, the ecological or conservation importance of old individual animals has been neglected. However, emerging literature has shown that old individuals provide distinct biological, ecological, cultural, and ecosystem service functions that have been—or are being—lost as these aged individuals are exploited. In this Review, we highlight the possible global consequences of the loss of old and often wise individual animals in the wild and emphasize the importance of longevity conservation as a solution. We highlight examples to encourage others to address new research, policy, and management avenues to better understand and conserve the specific ecological roles and ecosystem services that are provided by older, often larger, and more experienced animals.

**ADVANCES:** A poorly appreciated shift toward younger age structures in animal populations is causing harmful ecological, social, and economic consequences. Research on diverse vertebrate and invertebrate taxa has demonstrated the importance of old individuals to fundamental biological and ecological processes, including reproduction and recruitment, information acquisition and cultural transmission, trophic dynamics, and population resistance and resilience to ecological and anthropogenic disturbances. The importance of old individuals spans taxa that range from crustaceans, mollusks, sponges, and corals to fishes, turtles, birds, elephants, whales, carnivores, and humans. The benefits of old age contrast with the senescence-focused paradigm of aging but are supported by life history theory, behavioral ecology, and old animal storage effects. Life history theory explains why old animals of long-lived species often have a disproportionate influence on reproduction and population stability, which in turn affects how we manage fisheries, wildlife, and biodiversity. The loss of old individuals not only erodes reproductive and knowledge-related storage effects of these taxa but, owing to their roles in ecosystems (e.g., food webs, nutrient cycling, and ecosystem engineering), the loss of old animals has important consequences for ecosystem structure, function, and resilience in the face of global change.

**OUTLOOK:** New research, policy, and longevity conservation actions are needed to protect and restore the distinct ecological roles and ecosystem services provided by older, often larger, and more experienced animals. Longevity overfishing should be recognized as a type of overfishing, and longevity conservation may be considered an obligation of the United Nations Convention on Biological Diversity to maintain ecosystem structure and function. Despite the loss of old individuals contributing to population decline, collapse, and extinction, the International Union for the Conservation of Nature (IUCN) Red List does not yet recognize truncated age structure in its criteria for listing threatened species. Longevity conservation should become an explicitly stated objective for local, national, and international management of certain species of fish and wildlife. This will require methods for detecting and measuring longevity depletion (i.e., age truncation) when formulating policies for sustainable fisheries, wildlife management, and the recovery of threatened species.

Longevity conservation to reduce mortality of old individuals is feasible using strategies such as age- and size-based harvest regulations, catch and release, time-area restrictions, and networks of protected areas. Research and management priorities should include understanding (i) how old animals may yield storage effects that enhance population resilience and community stability, especially with regards to exploited fish and wildlife; (ii) how older cohorts adapt to and recover from climate change and other anthropogenic pressures; (iii) trophic dynamics of old and large individuals and their role in food web structure and function; (iv) how knowledge, sociality, culture, decision-making, and other behaviors exhibited by older individuals influence social hierarchies, migration, habitat use, and survival; and (v) how to implement policies and strategies to restore and protect the age structure of longevity-depleted populations. ■

**African savannah elephant (*Loxodonta africana*) grandmother, newborn calf, and family.** Despite the negative aspects of biological aging, old individual animals provide vital ecological and ecosystem service functions, which require longevity conservation. [Photo credit: Phyllis C. Lee, Amboseli Trust for Elephants]

---

The list of author affiliations is available in the full article online.
*Corresponding author. Email: keller.kopf@cdu.edu.au

Cite this article as R. K. Kopf *et al.*, *Science* **387**, eado2705 (2025). DOI: 10.1126/science.ado2705

**READ THE FULL ARTICLE AT**
https://doi.org/10.1126/science.ado2705

RESEARCH ARTICLE SUMMARY

CANCER GENOMICS

# Rare germline structural variants increase risk for pediatric solid tumors

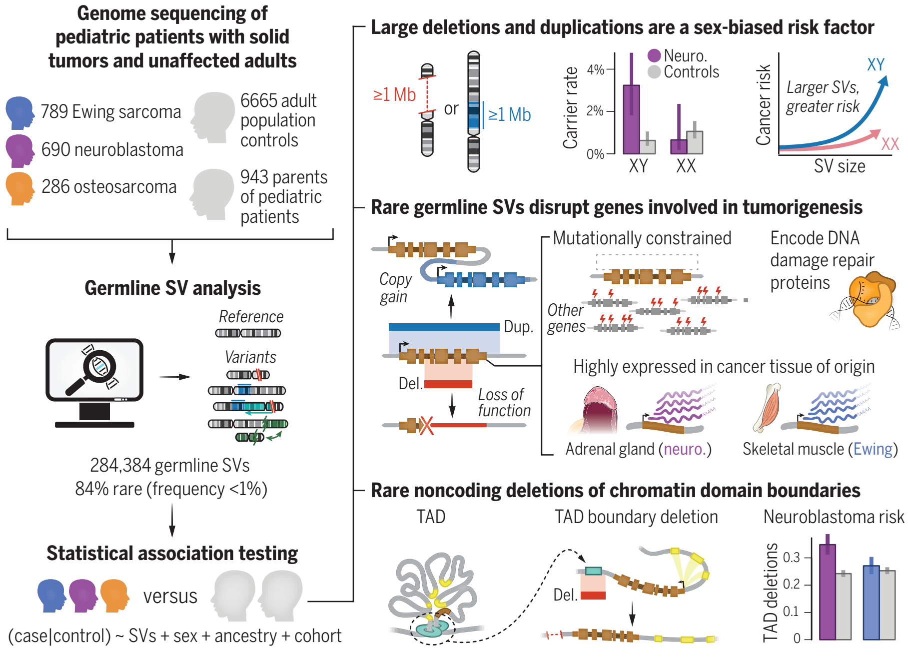
**Riaz Gillani*†, Ryan L. Collins†, Jett Crowdis, Amanda Garza, Jill K. Jones, Mark Walker, Alba Sanchis-Juan, Christopher W. Whelan, Emma Pierce-Hoffman, Michael E. Talkowski, Harrison Brand, Kevin Haigis, Jaclyn LoPiccolo, Saud H. AlDubayan, Alexander Gusev, Brian D. Crompton, Katherine A. Janeway, Eliezer M. Van Allen***

**INTRODUCTION:** Extracranial solid tumors make up one-third of all new pediatric cancer diagnoses annually and are a leading cause of childhood morbidity and mortality. The somatic drivers of these cancers have been meticulously characterized by prior studies, but the biomedical community lacks a comparable understanding of the earliest risk factors that lead to tumor development. Germline genetic factors are thought to play an important role in disease pathogenesis owing to the early age of onset and lower tumor mutational burden in these cancers, but presently recognized predisposition genes only account for 10 to 15% of all diagnoses. Identifying more germline risk factors in pediatric solid tumors might unlock a deeper understanding of the central tumorigenic biology and ultimately contribute to more informed approaches to clinical management and treatment.

**RATIONALE:** Most prior germline genetic studies in pediatric solid tumors have focused on short variants, specifically single-nucleotide variants and small insertions or deletions in the protein-coding regions of focused sets of genes. Structural variants (SVs) are a relatively understudied class of germline variation that involve rearrangements of ≥50 nucleotides. We hypothesized that rare germline SVs contribute to pediatric solid tumor predisposition and therefore undertook a study to evaluate their impact across coding and noncoding genomic regions in neuroblastoma, Ewing sarcoma, and osteosarcoma. We analyzed germline genome sequencing from 1765 affected patients, their 943 unaffected parents, and 6665 adult controls and statistically compared patterns of rare germline SVs in pediatric patients versus adult controls.

**RESULTS:** We identified a median of 7275 germline SVs per genome. Overall, de novo germline SV rates in neuroblastoma and Ewing sarcoma trios were comparable to those of the general population (~0.2 per genome), and de novo SVs were significantly larger and more likely to be gene-disruptive than inherited SVs. We discovered that ultrarare germline deletions or duplications of more than 1 million nucleotides were a sex-biased risk factor for pediatric solid tumors, exhibiting a fourfold enrichment in male patients relative to cancer-free controls. By implementing a category-wide association study framework, we found that germline SVs in mutationally constrained genes were enriched in neuroblastoma patients, and SVs that affect DNA damage repair genes were enriched in both Ewing sarcoma and neuroblastoma. Cross-examination of these gene-disruptive SVs with gene expression data revealed that the SVs in patients tended to disrupt genes that are highly expressed in their cancer's tissue of origin, and the dysregulated expression caused by these SVs manifested in the patients' tumors. Across our cohort, we also identified dozens of germline SVs that affect established cancer predisposition genes, such as *PHOX2B* and *BARD1* in neuroblastoma and Fanconi anemia genes in Ewing sarcoma, demonstrating that germline SV analyses can uncover new clinically relevant findings in individual patient genomes. Finally, we discovered that ultrarare SVs that overlap chromatin domain boundaries were enriched in neuroblastoma versus adult controls, which was not the case for Ewing sarcoma. This implied that subsets of rare noncoding SVs outside of protein-coding genes likely contribute to the pathogenesis of some pediatric solid tumors.

**CONCLUSION:** Overall, our results demonstrate that rare germline SVs that span coding and noncoding loci are important risk factors for pediatric solid tumors. We observed stronger SV effects in neuroblastoma relative to Ewing sarcoma, which is consistent with the markedly earlier age of onset in neuroblastoma. We estimate that rare SVs contribute 1.1 to 5.6% of liability for these cancers, which underscores a modest yet meaningful role for rare germline SVs in the missing heritability of pediatric solid tumors. Our work supports the integration of germline SVs into clinical practice and research efforts related to pediatric solid tumor predisposition. ■



**Rare germline SVs are risk factors for pediatric solid tumors.** We analyzed germline genome sequencing from 1765 patients with solid tumors, their 943 unaffected parents, and 6665 adult controls. Focusing on rare SVs, we found that large events, events that disrupted genes from biologically relevant pathways, and events that affected chromatin domain boundaries, were risk factors for pediatric solid tumors. Del., deletion; Dup., duplication; neuro., neuroblastoma; TAD, topologically associated domain.

---

The list of author affiliations is available in the full article online.
*Corresponding author. Email: eliezerm_vanallen@dfci.harvard.edu (E.M.V.A.); riaz_gillani@dfci.harvard.edu (R.G.)
†These authors contributed equally to this work.

**Cite this article as** R. Gillani *et al.*, *Science* **387**, eadq0071 (2025). DOI: 10.1126/science.adq0071

**READ THE FULL ARTICLE AT**
https://doi.org/10.1126/science.adq0071

RESEARCH ARTICLE SUMMARY

PLANT GENETICS

# Ancient structural variants control sex-specific flowering time morphs in walnuts and hickories

**Jeffrey S. Groh*, Diane C. Vik, Matthew Davis, J. Grey Monroe, Kristian A. Stevens, Patrick J. Brown, Charles H. Langley, Graham Coop***

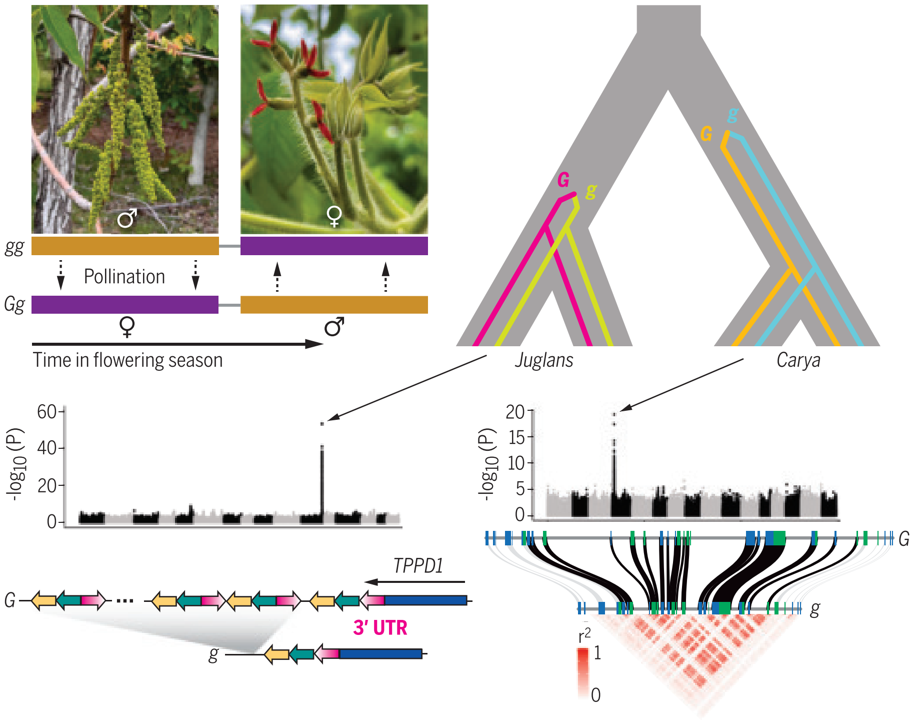
**INTRODUCTION:** Despite the capacity of many hermaphroditic plants to self-fertilize, the genetic costs of inbreeding have repeatedly driven the evolution of outcrossing strategies such as dioecy. Dioecy and other discrete mating type systems are maintained by rare mating type advantage and are often controlled by sets of tightly linked epistatic alleles (i.e., supergenes). We investigate the genetic underpinnings and evolution of a dimorphic mating system that exploits the temporal dimension - heterodichogamy. In species of hermaphroditic walnuts (*Juglans*) and hickories (e.g., pecans, *Carya*), two morphs alternate staminate and pistillate (i.e. male and female) flowering phases across the season, promoting disassortative mating that maintains a stable equilibrium. The genetic underpinnings and evolution of this mating system have remained elusive.

**RATIONALE:** Heterodichogamy is known to be inherited through a single locus system in both walnuts and hickories (Juglandaceae), with a dominant allele for female-first flowering (*G*, dominant; *g*, recessive). These patterns suggested that heterodichogamy could have evolved in their common ancestor. We investigated the molecular basis of the inheritance of heterodichogamy in these genera, its evolutionary origins and maintenance over deep time scales, and its genetic parallels to dioecy.

**RESULTS:** We identified a narrow (~20-kb) genomic region controlling flowering types in Northern California black walnut (*Juglans hindsii*), and further found this region to be associated with flowering type across the entire walnut genus, spanning ~40 million years of evolutionary history. This region harbors a clear candidate gene (*TPPD1*) linked to trehalose-6-phosphate metabolism, a known flowering time pathway. Patterns of shared polymorphism across *Juglans* localize to the 3′ untranslated region (UTR) of *TPPD1* and an adjacent tandem array of 8 to 12 inverted 3′ UTR paralogs found only in the dominant haplotype. Presence of the dominant haplotype is associated with tissue-specific increased *cis* expression of *TPPD1* in early development of male flowers, predicting their delayed maturity. We identified a set of small RNAs transcribed from this tandem array at the same developmental stage and suggest that these may play a role in regulating TPPD1.

In pecan (*Carya illinoinensis*), we identified a separate 200 to 445-kb region controlling heterodichogamy. This region shows strong linkage disequilibrium spanning ~20 protein-coding genes, several of which are linked to flower development. These haplotypes segregate throughout the *Carya* genus, and we estimate they diverged >50 million years ago. A lack of recombination in the near-permanently heterozygous dominant haplotype is accompanied by reduced genetic diversity and an accumulation of repetitive content, mirroring Y chromosome–like evolution. Neither the *Juglans* nor *Carya* locus segregates in other heterodichogamous genera within Juglandaceae, suggesting the existence of additional mechanisms.

**CONCLUSION:** Heterodichogamy can evolve through simple inheritance mechanisms that co-opt existing regulatory pathways and genetic architecture. These two distinct mechanisms in *Juglans* and *Carya* are deeply conserved but reflect either convergent evolution of the mating system or turnover of its control from an ancestral genetic system. Although these mechanisms are parallel to the genetics of dioecy, conservation of hermaphroditic function in heterodichogamy may limit the evolutionary advantage of recombination suppression thought to underlie sex chromosome heteromorphism. In summary, genetic mechanisms for heterodichogamy showcase both dynamic evolution and considerable stability. ■



**Multiple genetic mechanisms control a balanced polymorphism for flowering order in walnuts and hickories.** In *Juglans* (top left) and *Carya*, two morphs show complementary temporal separation between male and female flowering (heterodichogamy). Mating types are controlled by two nonhomologous single-locus mechanisms that arose in the common ancestor of each genus, respectively. (Bottom left) Simplified schematic of a putative functional regulatory element at the *Juglans* locus. (Bottom right) Strong genotypic correlations across the *Carya* locus indicate a lack of recombination between two colinear haplotypes with similar gene content.

---

The list of author affiliations is available in the full article online.

*Corresponding author. Email: jgroh@ucdavis.edu (J.S.G.); gmcoop@ucdavis.edu (G.S.C.)

Cite this article as J. S. Groh *et al.*, *Science* **387**, eado5578 (2025). DOI: 10.1126/science.ado5578

**READ THE FULL ARTICLE AT**
https://doi.org/10.1126/science.ado5578

RESEARCH ARTICLE SUMMARY

BIOCHEMISTRY

# Landscape profiling of PET depolymerases using a natural sequence cluster framework

Hogyun Seo†, Hwaseok Hong†, Jiyoung Park†, Seul Hoo Lee, Dongwoo Ki, Aejin Ryu,
Hye-Young Sagong, Kyung-Jin Kim*

**INTRODUCTION:** Numerous PET-degrading enzymes (PETases) have been discovered in nature, including PETases from *Ideonella sakaiensis* (IsPETase) and leaf-compost cutinase (LCC), which are undergoing engineering for industrial applications. However, little is known about how many enzymes possess this capability or which enzymes have greater potential for degrading plastics. The natural diversity accessible through sequence databases may hold the key to addressing this issue, but so far there has been a lack of available approaches for identifying the fitness of promiscuous enzymes from sequence data.

**RATIONALE:** We have attempted to determine the distribution of fitness of a large set of enzymes, providing a comparative and holistic view of the evolutionary success of naturally occurring PETases. To organize the landscape of the polyesterase-lipase-cutinase family, we generated a two-dimensional semantic network with sequence homology distances analyzed by a neighborhood analysis module, which controls the network stringency by using distance histogram data for each sequence. On the basis of the connected components (clusters) and density distribution of nodes on the network, representative nodes were sampled with stratified and cluster sampling methods to experimentally measure fitness for PET-degrading activity and thermal stability.

**RESULTS:** The landscaping approach generated 170 clusters comprising 1894 nodes (2064 sequences) of the polyesterase-lipase-cutinase family, covering the majority of known PETases. The primary screening for PETase, using stratified sampling, showed that more than half of the selected nodes have detectable PET-degrading activity, with a few nodes possessing high fitness. The secondary screening for PETase, using cluster sampling, distinguished three clusters (C3, C25, and C158) containing nodes with high fitness, demonstrating the representativeness of the selected nodes by their positions. Out of the 158 nodes that participated in the screening rounds, N1319 (Mipa-P) from C158 and N1501 (Kubu-P) from C25 were identified as the most promising enzymes, exhibiting superior features in PET decomposition speed and durability compared with other known benchmarks. Further engineering with rationally selected cross-template mutations generated the Mipa-P$^{M19}$ and Kubu-P$^{M12}$ variants. The engineered Kubu-P$^{M12}$ outperformed the engineered benchmarks under high substrate loads in industrial conditions. Kubu-P$^{M12}$ even endured ethylene glycol as the solvent and reactant in enzyme-catalyzed PET glycolysis, yielding bis(2-hydroxyethyl) terephthalate as the main product at concentrations reaching tens of millimolar.

**CONCLUSION:** The landscaping-and-sampling framework developed in this study provides an effective method for exploring the fitness landscape of promiscuous protein families, facilitating the accumulation of sequence-to-activity data. This study identifies 90 previously unreported sequences with potential PETase activity, of which approximately 90% represent novel groups. These findings provide a broad overview of the distribution of the PETases within the constructed library, enabling relative fitness comparisons. Consequently, this analysis led to the identification of three unexplored protein lineages exhibiting high fitness for PETase activity and to identification of the most potent proteins, Mipa-P and Kubu-P. Notably, the Kubu-P$^{M12}$ variant engineered from a high potential template outperforms benchmarks under industrially relevant conditions. ■

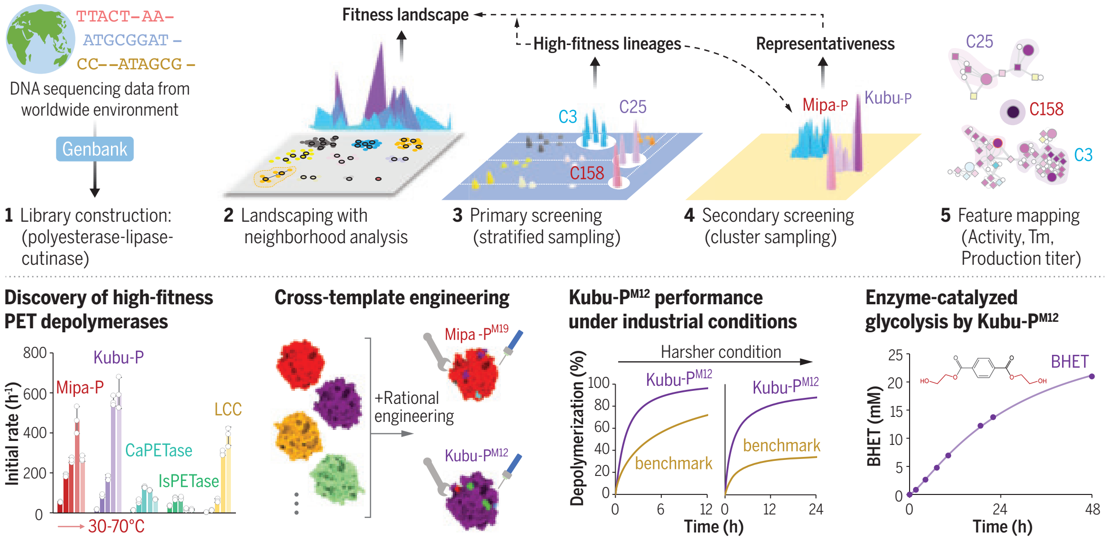
---

The list of author affiliations is available in the full article online.
***Corresponding author. Email: kkim@knu.ac.kr**
†These authors contributed equally to this work.

**Cite this article as H. Seo *et al.*, *Science* 387, eadp5637 (2025). DOI: 10.1126/science.adp5637**

**READ THE FULL ARTICLE AT**
https://doi.org/10.1126/science.adp5637

**Landscape profiling for the development of a superior PET depolymerase.** (Top) A research flowchart is shown, indicating steps (in numerical order) for using a sequence database to assess the fitness landscape of naturally occurring PET-degrading enzymes in the polyesterase-lipase-cutinase family. (Bottom) The development of the most efficient PET depolymerase through this landscaping is summarized (from left to right) in the four panels.



Comparisons of PET hydrolysis speed and duration

Rational engineering leveraging the diversity of templates in the landscape

Kubu-P$^{M12}$ from the high-fitness template outperforms benchmarks

Kubu-P$^{M12}$ enables significant levels of enzyme-catalyzed PET glycolysis

RESEARCH ARTICLE SUMMARY

MALARIA

# Protective antibodies target cryptic epitope unmasked by cleavage of malaria sporozoite protein

**Cherrelle Dacon†, Re'em Moskovitz†, Kristian Swearingen, Lais Da Silva Pereira, Yevel Flores-Garcia, Maya Aleshnick, Sachie Kanatani, Barbara Flynn, Alvaro Molina-Cruz, Kurt Wollenberg, Maria Traver, Payton Kirtley, Lauren Purser, Marlon Dillon, Brian Bonilla, Adriano Franco, Samantha Petros, Jake Kritzberg, Courtney Tucker, Gonzalo Gonzalez Paez, Priya Gupta, Melanie J. Shears, Joseph Pazzi, Joshua M. Edgar, Andy A. Teng, Arnel Belmonte, Kyosuke Oda, Safiatou Doumbo, Ludmila Krymskaya, Jeff Skinner, Shanping Li, Suman Ghosal, Kassoum Kayentao, Aissata Ongoiba, Ashley Vaughan, Joseph J. Campo, Boubacar Traore, Carolina Barillas-Mury, Wathsala Wijayalath, Azza Idris, Peter D. Crompton, Photini Sinnis, Brandon K. Wilder, Fidel Zavala, Robert A. Seder, Ian A. Wilson, Joshua Tan***

**INTRODUCTION:** Malaria is a major cause of global morbidity and mortality. Two World Health Organization–recommended malaria vaccines, RTS,S/AS01 and R21/Matrix M, are being deployed in endemic regions, but new interventions are still urgently needed to reduce malaria burden. Both recommended vaccines and the monoclonal antibodies (mAbs) being developed as clinical products target the central repeat region or closely related epitopes within the *Plasmodium falciparum* (*Pf*) circumsporozoite protein (PfCSP), the major sporozoite coat protein. Although PfCSP is a critical sporozoite protein, much remains to be discovered about the broader antigenic landscape of the sporozoite surface. An exploration of the human antibody response to this surface after malaria infection or vaccination could further our understanding of sporozoite biology and reveal targets that can be used to complement existing interventions.

**RATIONALE:** We developed an antigen-agnostic pipeline to survey single human B cells for the production of antibodies that bind to the surface of intact *Pf* sporozoites. To increase the probability of identifying unknown targets, we performed several deselection steps to exclude antibodies that target conventional PfCSP epitopes. The isolated sporozoite-binding mAbs were then used as tools to identify their target epitopes and investigated for in vitro and in vivo function against the malaria parasite.

**RESULTS:** We screened plasma from 941 *Pf* sporozoite–exposed individuals and identified five donors who retained antibody reactivity to *Pf* sporozoites after preblocking with recombinant PfCSP. Using our antigen-agnostic workflow, we identified 10 mAbs from these donors that bound strongly to *Pf* sporozoites but not to recombinant PfCSP. Through a series of biochemical and cellular assays, we found to our surprise that these mAbs bound exclusively to sporozoite-expressed PfCSP, because they required two sequential parasite-driven modifications to this protein: cleavage of the N terminus of PfCSP and conversion of the resultant N-terminal residue from glutamine to pyroglutamate. Through mass spectrometric analysis, we found that this epitope, which we named pGlu-CSP, is commonly found on the sporozoite surface. X-ray structural studies confirmed that the cyclic pyroglutamate residue, along with adjacent amino acids specific to PfCSP, are critical binding residues that fit within a pocket formed within the pGlu-CSP–reactive mAbs. The minimal epitope bound by these mAbs is pGlu$^{96}$PADGNP$^{102}$, which is distinct from the target epitopes of previously isolated PfCSP-reactive mAbs. The mAb MAD21-101 conferred sterile protection against *Pf* sporozoite infection by mosquito bite in the FRG-HuHep human liver-chimeric mouse model. It did not bind to the malaria vaccine R21 (which uses the same PfCSP sequence as RTS,S), indicating that it is unlikely to interfere with vaccine function if used in regions where the vaccines are deployed.

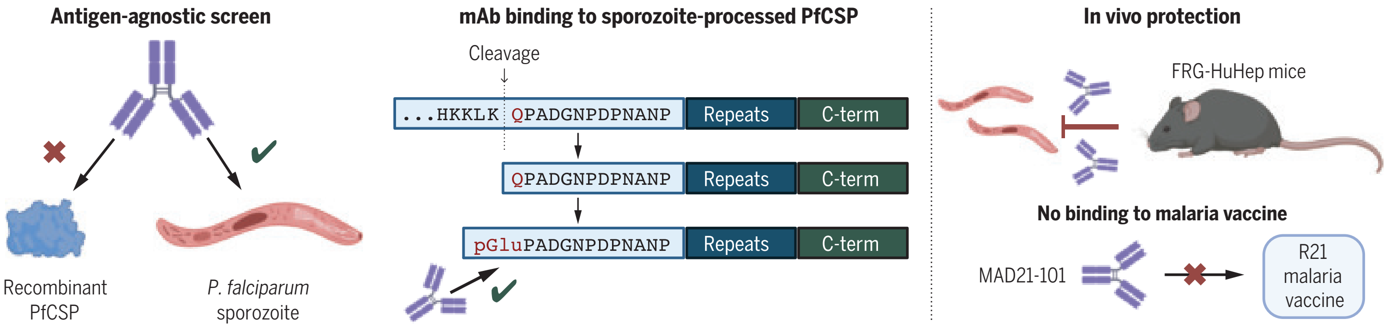
**CONCLUSION:** This study identified the cryptic pGlu-CSP site on the *Pf* sporozoite surface as a target of protective antibodies. To our knowledge, these are the first potent antisporozoite mAbs that do not bind to the PfCSP central repeat region, which is included in both RTS,S/AS01 and R21/Matrix M. Therefore, they are unlikely to interfere with vaccines that are already being deployed, providing a rationale for clinical development. The identification of these mAbs also provides a resource for the research community to further investigate PfCSP cleavage by sporozoites, which has been associated with hepatocyte invasion but is still the subject of active investigation. More broadly, these findings show the utility of using an antigen-agnostic approach in identifying targets of protective immunity against infectious pathogens. ■

---

The list of author affiliations is available in the full article online.
***Corresponding author. Email: tanj4@nih.gov**
†These authors contributed equally to this work.

**Cite this article as C. Dacon *et al.*, *Science* 387, eadr0510 (2025). DOI: 10.1126/science.adr0510**

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adr0510



**Antibodies targeting the sporozoite pGlu-CSP site protect against malaria.** Left: an antigen-agnostic screen led to the discovery of antibodies that bind *Pf* sporozoites but not the recombinant form of the major coat protein, PfCSP. These antibodies bound exclusively to sporozoite-expressed PfCSP because they required two sequential parasite-mediated modifications. Right: the mAb MAD21-101 provides sterile protection in an in vivo model of *Pf* sporozoite invasion and does not bind to the R21 malaria vaccine. [Figure created with BioRender.com]

RESEARCH ARTICLES

PHOTORECEPTORS

# Direct photoreception by pituitary endocrine cells regulates hormone release and pigmentation

Ayaka Fukuda[1]†, Keita Sato[2], Chika Fujimori[1]‡, Takahiro Yamashita[3], Atsuko Takeuchi[4], Hideyo Ohuchi[2], Chie Umatani[5]§, Shinji Kanda[1]*

**The recent discovery of nonvisual photoreceptors in various organs has raised expectations for uncovering their roles and underlying mechanisms. In this work, we identified a previously unrecognized hormone-releasing mechanism in the pituitary of the Japanese rice fish (medaka) induced by light. $Ca^{2+}$ imaging analysis revealed that melanotrophs, a type of pituitary endocrine cell that secretes melanocyte-stimulating hormone, robustly increase the concentration of intracellular $Ca^{2+}$ during short-wavelength light exposure. Moreover, we identified Opn5m as the key molecule that drives this response. Knocking out *opn5m* attenuated melanogenesis by reducing *tyrosinase* expression in the skin. Our findings suggest a mechanism in which direct reception of short-wavelength light by pituitary melanotrophs triggers a pathway that might contribute to protection from ultraviolet radiation in medaka.**

A growing body of evidence shows that various organs, including the central nervous system, express nonvisual opsin genes that may contribute to the perception of light information in vertebrates (*1*–*4*). For instance, nonvisual opsins in the brain have been reported to regulate seasonal breeding in Japanese quail and to modulate thermogenesis in mice (*5*, *6*). In addition, there is an intriguing evolutionary aspect that nonmammalian vertebrates directly perceive light in the pineal gland to modulate circadian rhythms (*7*, *8*), whereas mammals have lost this mechanism and modulate their circadian rhythms by light information from the eyes via a compensatory neural pathway (*9*). Here, using a nonmammalian model teleost, the Japanese rice fish medaka (*10*–*13*), we identified a pathway in the pituitary, located at the base of the brain, by which the endocrine cells, melanotrophs, directly receive light and autonomously secrete hormones. We further investigated the mechanism and biological function of this phenomenon.

## Melanotrophs of the pituitary directly respond to short-wavelength light, which leads to increased $[Ca^{2+}]_i$

First, we performed $Ca^{2+}$ imaging of melanotrophs using a whole brain–pituitary in vitro preparation in medaka. We used *pro-opiomelanocortin a* (*pomca*):GCaMP animals, in which, under the *cis*-regulatory activity of the 3.7-kb 5′ flanking region of *pomca*, melanotrophs and corticotrophs specifically express a $Ca^{2+}$ indicator, GCaMP6s (fig. S1) (*14*). During $Ca^{2+}$ imaging, a robust increase in the concentration of intracellular $Ca^{2+}$ ($[Ca^{2+}]_i$) in melanotrophs was observed without any stimulation except exposure to the blue excitation light for fluorescence observation. This phenomenon suggests that some regions in the brain or pituitary received light and induced the $[Ca^{2+}]_i$ increase in the melanotrophs (fig. S2). To determine whether light is sensed by the pituitary or the brain, we performed $Ca^{2+}$ imaging using an isolated pituitary. An increase in $[Ca^{2+}]_i$ was observed in melanotrophs within a few seconds of blue excitation light exposure, even in the isolated pituitary (Fig. 1, A and B, and movie S1). We observed this phenomenon in pituitary from both male and female animals (Fig. 1, A and B for males and fig. S3 for females). On the contrary, corticotrophs, which express adrenocorticotropic hormone (ACTH) (*15*) derived from the same precursor, POMC, did not exhibit a $[Ca^{2+}]_i$ increase under the same condition (Fig. 1, A and B). Because the robust rise in $[Ca^{2+}]_i$ triggers hormone release in endocrine cells (*16*–*18*), we hypothesized that isolated pituitary could autonomously release melanocyte-stimulating hormone (MSH) by directly sensing light. Next, we analyzed the wavelength dependence of this nonvisual photoreception of the pituitary using stimulation light with different wavelengths. To avoid the unintentional direct detection of stimulation light during image acquisition, we delivered light stimulation only during the interval of each imaging acquisition using an Arduino (microcontroller)–based control system (fig. S4). We evaluated the effects of light stimulation (365 to 740 nm, ~200 μmol $m^{-2}$ $s^{-1}$) applied during recording intervals (excitation exposure, 50 ms; recording interval, 5 s). Because the data indicated that melanotrophs exhibit a steady $[Ca^{2+}]_i$ after 200 s under this imaging condition (fig. S5A), we began the evaluation at least 200 s after the 5-s interval fluorescence image acquisition had started. The time course of this experiment is detailed in fig. S5B. We found that melanotrophs respond strongly to short-wavelength light, especially ultraviolet (UV) light (Fig. 1, C and D, and fig. S5C), which suggests the involvement of short wavelength–sensitive photoreceptors in this mechanism. Next, to examine whether a single melanotroph directly receives light or whether other photoreceptive cells in the pituitary mediate this light-induced $[Ca^{2+}]_i$ increase, we performed $Ca^{2+}$ imaging of dissociated cells from the posterior part of the pituitary of *pomca*:GCaMP medaka, where melanotrophs are localized. Isolated melanotrophs labeled by GCaMP responded to the excitation light (Fig. 1E), which is similar to the response observed in the whole pituitary preparation. As expected, three repeated trials resulted in a $[Ca^{2+}]_i$ increase in an isolated cell (fig. S6). These results demonstrate that melanotrophs, a type of endocrine cell, receive short-wavelength light and respond by increasing $[Ca^{2+}]_i$. Additionally, we explored whether the $Ca^{2+}$ source responsible for the light-induced $[Ca^{2+}]_i$ increase was extracellular fluid or intracellular store. Perfusion of $Ca^{2+}$-free artificial cerebrospinal fluid (ACSF) or ACSF that contained an inhibitor of $Ca^{2+}$ channels, $Cd^{2+}$ ($CdCl_2$), did not affect light-induced $[Ca^{2+}]_i$ increase. By contrast, perfusion of ACSF that contained an inhibitor of $Ca^{2+}$ release from endoplasmic reticulum (ER), 2-aminoethoxydiphenyl borate, substantially diminished light-induced $[Ca^{2+}]_i$ increase (Fig. 1F and fig. S7), which suggests that $Ca^{2+}$ is derived from intracellular ER. Given that melanotrophs are endocrine cells, these results suggest the presence of an autonomous release of hormone in response to light.

## Opn5m is the crucial molecule for photoreception of melanotrophs

Next, we investigated the mechanism involved in direct photoreception in melanotrophs by analyzing which photoreceptor could be responsible for the observed effects. Among the 34 medaka opsins, in situ hybridization indicated that only *opn5m* is expressed in the posterior part of the pituitary, where melanotrophs are localized (figs. S8 and S9) (*19*). Orthologs

---

[1]Department of Marine Bioscience, Atmosphere and Ocean Research Institute, The University of Tokyo, Kashiwa, Japan. [2]Department of Cytology and Histology, Faculty of Medicine, Dentistry, and Pharmaceutical Sciences, Okayama University, Okayama, Japan. [3]Department of Biophysics, Graduate School of Science, Kyoto University, Kyoto, Japan. [4]Division of Analytical Laboratory, Kobe Pharmaceutical University, Kobe, Japan. [5]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, Japan.
***Corresponding author. Email: shinji@aori.u-tokyo.ac.jp**
†Present address: Division of Behavioral Neurobiology, National Institute for Basic Biology, Okazaki, Japan.
‡Present address: Department of Biological Sciences, School of Science, Hokkaido University, Sapporo, Japan.
§Present address: Graduate School of Agriculture, Tokyo University of Agriculture and Technology, Tokyo, Japan.

of *opn5m*, also referred to as neuropsin, have been shown to be expressed in the hypothalamic region of other animals (*5*, *6*, *20*). Using double in situ hybridization in conjunction with *pomca*, we found co-expression of *opn5m* and *pomca* in the posterior part of the pituitary, which indicates that melanotrophs express *opn5m* mRNA (Fig. 2A and fig. S10). Additionally, we analyzed the UV-visible absorbance spectrum of recombinant medaka Opn5m. The 11-*cis*-retinal-bound resting state of medaka Opn5m showed maximum absorption in the short-wavelength region (Fig. 2B and fig. S11), which is consistent with the properties of other members of the Opn5m group, such as zebrafish Opn5m (*21*, *22*). These findings are in agreement with the $Ca^{2+}$ imaging result that melanotrophs responded strongly to short-wavelength light (Fig. 1D). To support the hypothesis that Opn5m is the key molecule responsible for light sensitivity in melanotrophs, we disrupted the *opn5m* gene in *pomca*:GCaMP medaka using CRISPR/Cas9 (fig. S12) (*23*). The mutant *opn5m* gene encodes a truncated protein lacking transmembrane helices 5, 6, and 7, including the putative retinal-binding-site K294, which is predicted to be a loss-of-function mutation (*2*). In $Ca^{2+}$ imaging, melanotrophs of *opn5m*$^{-/-}$ medaka did not respond to the excitation light, whereas those of *opn5m*$^{+/-}$ responded similarly to the wild type (WT) (Fig. 2, C to E). These results indicate that Opn5m is indispensable for the short-wavelength-light–induced $[Ca^{2+}]_i$ increase in melanotrophs. To confirm that this defect is specific to response to light, we examined the effect of corticotropin-releasing hormone (CRH), which is known to induce the release of MSH in another teleost (*24*). The receptor of CRH, *crhr1*, is expressed in the posterior part of the medaka pituitary, where melanotrophs reside (fig. S13), and our data show that CRH increased the $[Ca^{2+}]_i$ in both *opn5m*$^{+/-}$ and *opn5m*$^{-/-}$ animals, which shows that the ablation of *opn5m* does not affect the response to CRH in melanotrophs (Fig. 2F and fig. S14). We also found that UV light stimulation has an additive effect on the increase in $[Ca^{2+}]_i$ in conjunction with CRH (fig. S15). Furthermore, we tested whether Opn5m was sufficient for light-induced $[Ca^{2+}]_i$ increase by conducting $Ca^{2+}$ imaging in a widely used human cell line, human embryonic kidney (HEK) 293A cells, and a rodent pituitary cell line, LβT2 cells (*25*), transfected with *opn5m* and *gcamp6s*. As expected, after incubation with an opsin chromophore, 11-*cis*-retinal or all-*trans*-retinal, these Opn5m-expressing HEK cells and LβT2 cells showed $[Ca^{2+}]_i$ increases during blue excitation light exposure (Fig. 2, G and H, and fig. S16), which is consistent with previous studies of mammalian Opn5m (*6*, *26*, *27*). We conclude that Opn5m is necessary and sufficient for the light-induced $[Ca^{2+}]_i$ increase.

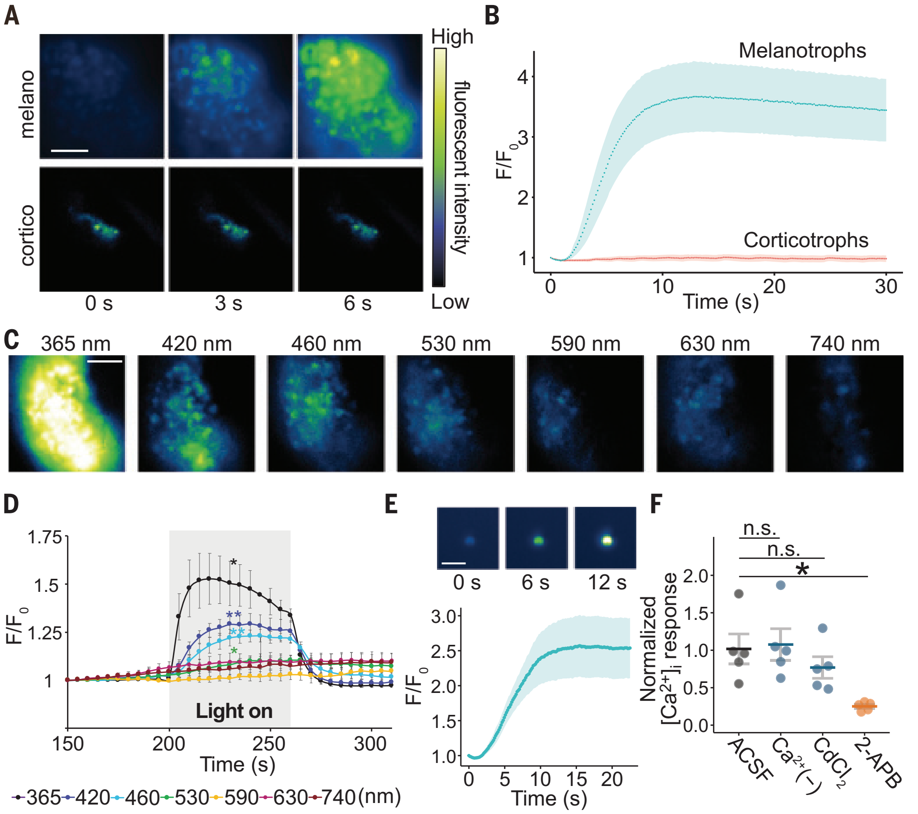


**Fig. 1. Short wavelength light–induced $[Ca^{2+}]_i$ rise of melanotrophs in the pituitary.** (**A**) Representative image series showing fluorescence changes during the excitation light exposure (450 to 490 nm) in melanotrophs (melano) and corticotrophs (cortico) of *pomca*:GCaMP medaka. (**B**) Time course of GCaMP fluorescence changes of melanotrophs and corticotrophs; *n* = 5 male medaka. (**C**) Representative GCaMP fluorescence images of melanotrophs, captured 40 s after exposure to the indicated wavelengths (365, 420, 460, 530, 590, 630, and 740 nm). All irradiated lights were adjusted to an ~200 (μmol $m^{-2}$ $s^{-1}$) integral sum of photon flux density. (**D**) Time course of GCaMP fluorescence changes in melanotrophs under stimulation with various wavelengths of light. Each wavelength was provided between 200 and 260 s after the recording began. *n* = 5 male medaka; $^{**}P$ = 0.0043 and 0.0057, $^{*}P$ = 0.012 and 0.016; paired *t* tests were applied to compare the average fluorescence for 50 s before and during irradiation for each wavelength. (**E**) Representative images of an isolated melanotroph labeled with GCaMP that showed increase in fluorescence during imaging with the blue excitation light (450 to 490 nm). Time course of GCaMP fluorescence changes of isolated melanotrophs; *n* = 5. (**F**) $Ca^{2+}$ response to the blue excitation light in melanotrophs during pharmacological inhibition by using $Ca^{2+}$-free ACSF $[Ca^{2+}(-)]$, 100 μM $CdCl_2$, and 500 μM 2-aminoethoxydiphenyl borate (2-APB). *n* = 5 medaka; $^{*}P$ = 0.011; n.s., not significant; Dunnet tests. Scale bars, 50 μm (A, C) and 25 μm (E). Data are represented as mean ± SEM.

### Direct photoreception of the pituitary can occur in natural conditions

Additionally, we elucidated the importance of direct photoreception by the pituitary even though the fish can acquire light information from the retina. We compared the $[Ca^{2+}]_i$ response in melanotrophs when white light-emitting diode (LED) light was applied to the retina or the pituitary, using a semi-intact preparation with the whole brain, pituitary, and eye covered by the skull except for the ventral surface of the pituitary (Fig. 3A and fig. S5C). This $Ca^{2+}$ imaging indicated that light stimulation to the pituitary, but not the retina, induces a $[Ca^{2+}]_i$ increase in melanotrophs (Fig. 3, B to D). Furthermore, we performed $Ca^{2+}$ imaging in larval medaka (5 to 8 days after hatching). Using an ex vivo preparation with an intact whole head, we found that the response to light in melanotrophs also occurred in early developmental stages (Fig. 3, E and F), which suggests that this mechanism can be applied to other, larger fishes in their larval stages, where their heads are likely more transparent because of thinner structures.

### MSH released by the photoreception of melanotrophs results in the enhancement of peripheral melanogenesis by increasing tyrosinase expression

To determine whether this light-induced $[Ca^{2+}]_i$ increase in melanotrophs induced hormone

**Fig. 2. Opn5m is necessary and sufficient for light sensitivity of melanotrophs.** (**A**) Double in situ hybridization that shows *opn5m* expression in melanotrophs. R, rostral; C, caudal. Scale bar, 50 μm. (**B**) Absorption spectrum of recombinant Opn5m. Red line indicates fitted spectra. (**C**) Representative image series that shows fluorescence changes during the excitation light exposure (450 to 490 nm) in melanotrophs of $opn5m^{+/-}$ or $opn5m^{-/-}$ medaka. Scale bar, 50 μm. (**D**) Time course of GCaMP fluorescence changes of melanotrophs in $opn5m^{+/-}$ or $opn5m^{-/-}$ medaka during exposure to the blue excitation light; $n$ = 5 male medaka. (**E** and **F**) Fold changes in GCaMP fluorescence of melanotrophs in response to light exposure (E) or CRH application (F) in $opn5m^{+/-}$ and $opn5m^{-/-}$ medaka. $n$ = 5 male medaka; *$P$ = 0.012; n.s., not significant; Student's $t$ test. (**G**) Representative image series that shows fluorescence changes during the excitation light exposure (450 to 490 nm) in HEK293A cells that express mock or *opn5m* and *gcamp6s*. Note that the cells were incubated with all-*trans*-retinal after transfection. Arrows indicate the cells transfected with *opn5m* and *gcamp6s*. Scale bar, 25 μm. (**H**) Maximum relative fluorescence of GCaMP during the whole recording (15 s) of each HEK293A cell that expressed GCaMP6s (mock, 133 cells) or cells that expressed Opn5m and GCaMP6s (175 cells). ***$P$ < 0.001; Student's $t$ test. Data are represented as mean ± SEM.

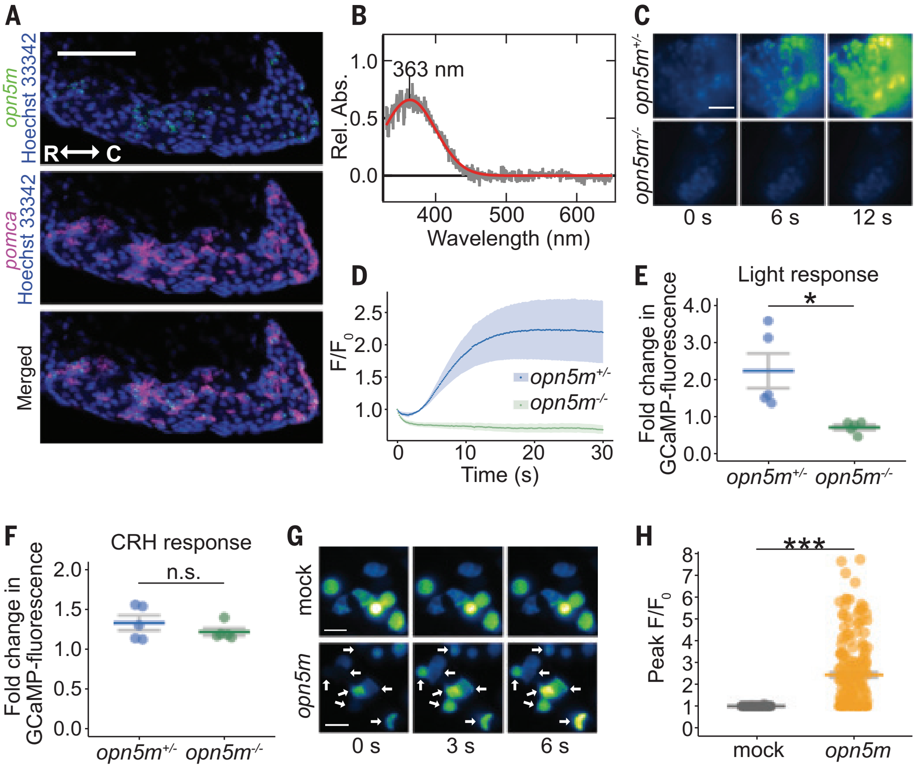
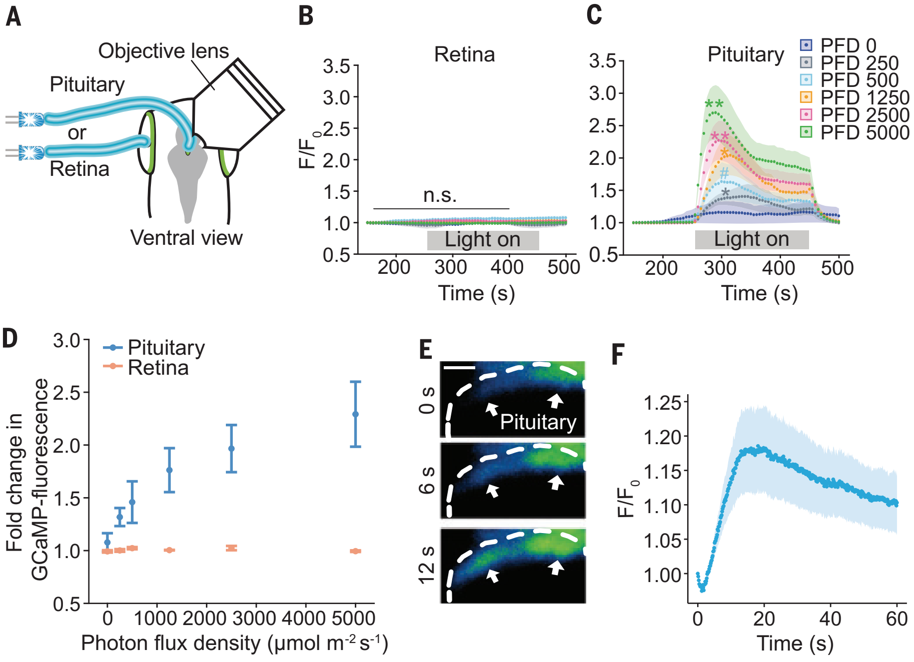
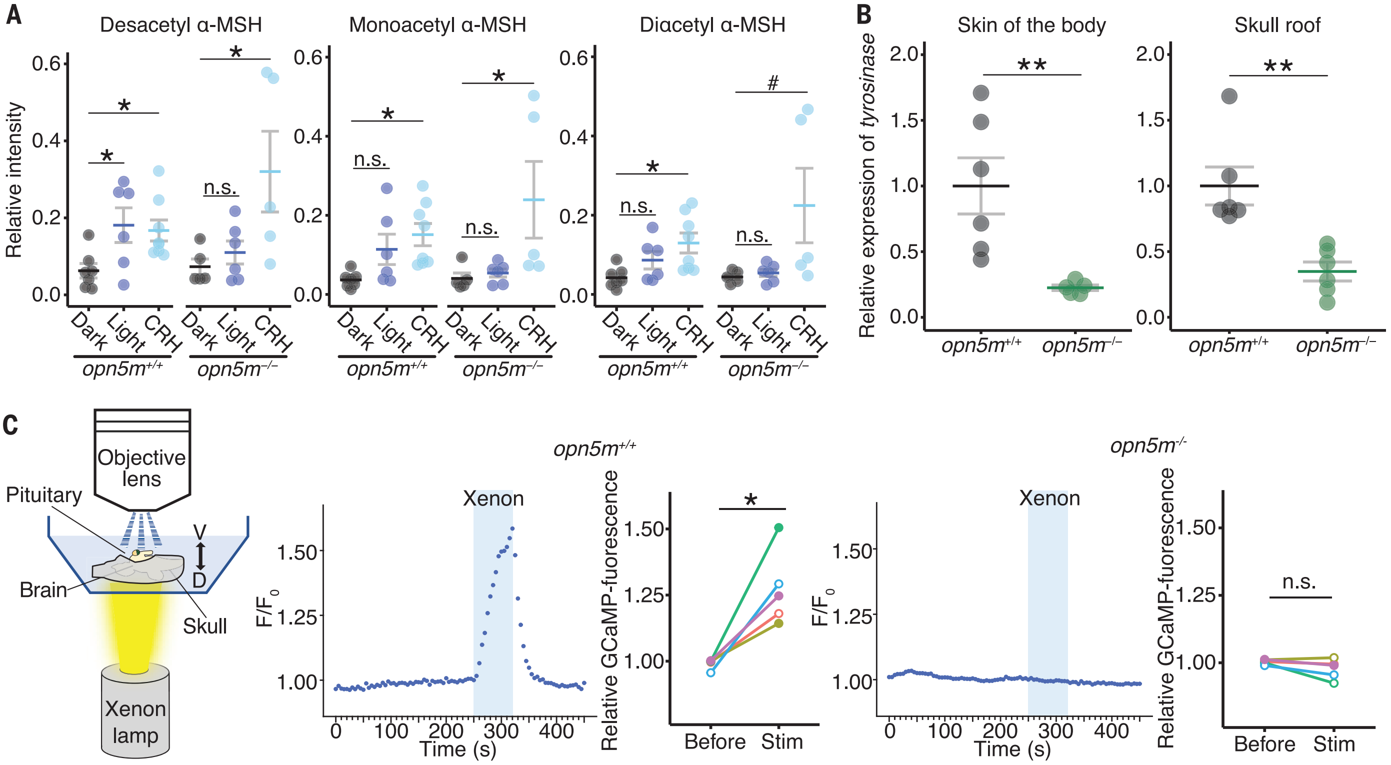


**Fig. 3. Direct photoreception in the pituitary is crucial for the $Ca^{2+}$ response of melanotrophs.** (**A**) Experimental scheme of $Ca^{2+}$ imaging with local light irradiation by using an optical fiber in semi-intact preparation (ventral view). The retina or pituitary was separately irradiated with white LED. (**B** and **C**) Time course of GCaMP fluorescence changes of melanotrophs with local light irradiation to (B) retina or (C) pituitary. $n$ = 7 male medaka; **$P$ = 0.0062 and 0.0049, *$P$ = 0.01, and #$P$ = 0.065, respectively; n.s., not significant; paired $t$ tests were applied before and during irradiation. PFD, photon flux density. (**D**) Effect of various intensities of white LED stimulation to pituitary (blue) or retina (orange). Fold changes of GCaMP fluorescence of melanotrophs are shown (during irradiation/before irradiation). (**E**) Representative image series that shows GCaMP fluorescence changes of melanotrophs during excitation light exposure (450 to 490 nm) in larva. Arrows indicate the melanotrophs. White dotted line indicates the pituitary region. Scale bar, 12.5 μm. (**F**) Time course of GCaMP fluorescence changes of melanotrophs in larvae; $n$ = 6 medaka. Data are represented as mean ± SEM.



release, we analyzed the contents of α-MSH derivatives in the medium in which isolated pituitaries were incubated. After applying 409-nm light or CRH for 3 hours, we analyzed the medium for α-MSH derivatives (desacetyl α-MSH, monoacetyl α-MSH, diacetyle α-MSH) using liquid chromatography–mass spectrometry analysis. We detected a light-dependent increase in desacetyl α-MSH peptide released from the pituitaries of WT medaka, but not from those derived from $opn5m^{-/-}$ animals. As expected, CRH application increased the release of desacetyl α-MSH from the pituitaries of both WT and $opn5m^{-/-}$ medaka (Fig. 4A). In the pituitaries of WT medaka, light stimulation did not affect the release of monoacetyl and diacetyl α-MSH compared with darkness, whereas CRH increased both monoacetyl and diacetyl α-MSH. This difference may be due to the distinct signaling pathways they activate: The CRH receptor activates $G_s$ proteins, which increases both intracellular cyclic adenosine 3′,5′-monophosphate (cAMP) (*28*) and $[Ca^{2+}]_i$ (figs. S14 and S15), whereas Opn5m activates



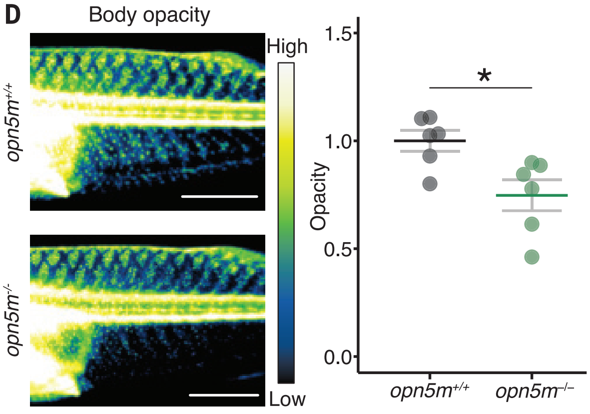


**Fig. 4. MSH is released by direct photoreception of melanotrophs, which results in increased *tyrosinase* expression in the skin.** (**A**) Liquid chromatography–mass spectrometry analysis of the pituitary culture supernatant, which were applied light or CRH stimulation. *n* = 5 to 8; *$P$ = 0.031, 0.041, 0.023, 0.011, 0.042, and 0.011 and #$P$ = 0.055, respectively; n.s., not significant; Dunnet tests. (**B**) Expression of *tyrosinase* in the skin and skull roof of *opn5m*$^{+/+}$ and *opn5m*$^{-/-}$ medaka. *n* = 6 or 5 male medaka; **$P$ = 0.0099 and 0.0023; Student's *t* test. (**C**) Left panel: experimental scheme of Ca$^{2+}$ imaging with xenon lamp stimulation in a preparation retaining all structures dorsal to the pituitary. Middle and right panels: representative time course of GCaMP fluorescence changes in melanotrophs with xenon lamp stimulation. Fold changes in GCaMP fluorescence are shown (middle: *opn5m*$^{+/+}$, right: *opn5m*$^{-/-}$). Before, average fluorescence from 0 to 30 s before stimulation; stimulation (Stim), average fluorescence from 40 to 70 s after stimulation began. *n* = 5 medaka; *$P$ = 0.012; n.s., not significant; paired *t* tests. Open circles indicate female medaka, and filled circles indicate male medaka. D, dorsal; V, ventral. (**D**) Left panel: body opacity of *opn5m*$^{+/+}$ and *opn5m*$^{-/-}$ exposed to mixed UV and white LED light for 3 days, shown by using pseudocolor. Right panel: quantification of body opacity in the dorsal area of *opn5m*$^{+/+}$ and *opn5m*$^{-/-}$ medaka. A wild strain, Kiyosu, was used. Scale bar, 2 mm. *n* = 6, 6 male medaka; *$P$ = 0.015; Student's *t* test. Data are represented as mean ± SEM.

$G_q$ proteins to increase $[Ca^{2+}]_i$ (*26*). The upregulation of MSH-related genes that occurred during the 3-hour incubation with CRH may have contributed to the observed difference. Because desacetyl α-MSH was released upon Opn5m activation, we examined its potential function in activating Mc1r, a key receptor in the melanogenesis pathway (*29*, *30*). Desacetyl α-MSH peptide activated the downstream $G_s$–adenylyl cyclase–cAMP pathway of Mc1r similarly to monoacetyl α-MSH, the major form of α-MSH in mammals (fig. S17). Although not all α-MSH derivatives showed a statistically significant increase—possibly owing to the small size of medaka and assay sensitivity limitations—our findings indicate that short-wavelength light induces the release of at least one α-MSH derivative that activates Mc1r. Taken together, the data indicate that short-wavelength light could induce exocytosis of melanotrophs through Opn5m.

Subsequently, we investigated the biological relevance of this UV-induced release of MSH at the whole-body level. Because it has been suggested that MSH increases the expression of *tyrosinase*, which is the rate-limiting enzyme in the melanogenesis pathway that is activated by MSH by means of Mc1r (*31*, *32*), we analyzed the expression of *tyrosinase* in *opn5m* knockout (KO) medaka. We exposed medaka to mixed LED light, including 365 nm and normal white LED, in a 14L:10D cycle for at least 3 days to simulate sunlight conditions (fig. S18A). The skin and the skull roof of *opn5m*$^{-/-}$ medaka showed a lower *tyrosinase* expression compared with *opn5m*$^{+/+}$ and *opn5m*$^{+/-}$ in both female and male medaka (Fig. 4B and figs. S19 and S20). Moreover, the expression of tyrosinase-related proteins 1a (*tyrp1a*) and 1b (*tryp1b*), which are also involved in the melanogenesis pathway, were lower in *opn5m*$^{-/-}$ medaka (figs. S18 to S20). We also confirmed that the downstream pathway of MSH is functional in *opn5m*$^{-/-}$ medaka by showing that an Mc1r agonist, [Nle$^4$,D-Phe$^7$]-α-MSH, increased the expression of *tyrosinase*, *tyrp1a*, and *tryp1b* (fig. S21). Thus, we conclude that the decrease in *tyrosinase* expression in *opn5m*$^{-/-}$ medaka is likely due to the attenuation of MSH release. In addition to the pituitary, in situ hybridization of the whole body in larvae showed that *opn5m* is also expressed in other organs, including some brain regions, the gut, the retina, and the skin (figs. S22 and S23). Because these results raised a possibility that peripheral photoreception could be the cause of the up-regulation of melanogenesis, we evaluated the role of pituitary photoreception in Opn5m-mediated *tyrosinase* expression. We used medaka with their pituitaries shaded by small black or transparent polyethylene sheets. After 1 day of exposure to mixed UV and white LED light, we examined the expression of *tyrosinase*, *tyrp1a*, and *tyrp1b* in the body skin. Medaka with black covers on their heads above the pituitary exhibited lower expression of *tyrosinase* and *tyrp1a* in the skin compared with those with transparent covers (fig. S24). These results suggest that the UV-induced hormone release from pituitary melanotrophs leads to up-regulation of *tyrosinase* and related protein expression in the skin, although this conclusion does not rule out the possibility of an additional pathway for UV-induced melanogenesis mediated by other tissues.

On the basis of the results so far, we hypothesized that the direct photoreception of melanotrophs may result in enhancement of melanogenesis under invasive UV light. Therefore, we next investigated the functional relevance of this elevated *tyrosinase* expression within a natural context. Because the model medaka strain used so far (himedaka) exhibits minimal black pigmentation in the body owing to posttranslational attenuation of tyrosinase activity, we additionally generated *opn5m* KO in a wild strain, Kiyosu, with normal black pigmentation (fig. S25) (*33*, *34*). First, we confirmed that *opn5m*$^{+/+}$ and *opn5m*$^{-/-}$ individuals of the Kiyosu strain exhibited a similar phenotype to those of the himedaka strain in both in vitro Ca$^{2+}$ imaging and in vivo *tyrosinase* expression properties (figs. S26 and S27). Next, to determine whether natural sunlight can induce $[\mathrm{Ca}^{2+}]_\mathrm{i}$ increase, we prepared a semi-intact preparation of the *pomca*:GCaMP Kiyosu strain that retained all structures above the pituitary (brain and skull roof) (Fig. 4C, left). We performed Ca$^{2+}$ imaging of melanotrophs and examined the effect of xenon lamp illumination, which has a spectrum similar to sunlight, from the dorsal side of the head. The light used was weaker than sunlight that reaches a depth of ~30 cm in water, where medaka typically live (fig. S28). Melanotrophs in *opn5m*$^{+/+}$ medaka showed a $[\mathrm{Ca}^{2+}]_\mathrm{i}$ increase in response to xenon lamp stimulation at an intensity comparable with that experienced under natural conditions (Fig. 4C, middle). As expected, melanotrophs in *opn5m*$^{-/-}$ medaka did not show a $[\mathrm{Ca}^{2+}]_\mathrm{i}$ increase in response to xenon lamp stimulation (Fig. 4C, right). Therefore, we conclude that the natural sunlight to which medaka are exposed is sufficient to elicit $[\mathrm{Ca}^{2+}]_\mathrm{i}$ increase in pituitary melanotrophs. In addition, we assessed the change in body opacity, which increases with melanogenesis. After 3 days of exposure to a mix of UV and white LED light, *opn5m*$^{-/-}$ Kiyosu medaka exhibited lower opacity in their distributed melanocytes and had lower body opacity in the dorsal area compared with their *opn5m*$^{+/+}$ siblings (Fig. 4D and fig. S29). This result is consistent with the observed reduction in *tyrosinase* and related genes in the skin of *opn5m* KO fish (Fig. 4B and figs. S18 to S20 for himedaka, fig. S27 for Kiyosu). Similar experiments with UV exposure for 4, 7, and 10 days also indicated that the body opacity of *opn5m*$^{-/-}$ medaka was lower than that of their *opn5m*$^{+/+}$ siblings (fig. S30). Taken together, these findings suggest that direct photoreception by pituitary melanotrophs can enhance the peripheral melanogenesis pathway when exposed to natural sunlight with strong UV light.

## Discussion

The present study identified a mechanism of the pituitary in which melanotrophs autonomously release hormones in response to short-wavelength light through Opn5m. Through this mechanism, MSH is released under strong UV environments (similar to sunlight), which enhances the melanogenesis pathway in the peripheral body surface (fig. S31). We speculate that this mechanism might contribute to the protection from invasive UV. In addition to the nonvisual opsins reported in various brain regions (*6*, *35*–*38*), this study demonstrates their physiological function in the pituitary, which is localized more deeply than the brain. This finding expands the potential roles of nonvisual opsins to a much wider range of organs than previously considered (*39*–*43*).

The present study successfully observed real-time cellular responses induced by nonvisual opsins with high temporal resolution through precise light stimulation and image acquisition. The application of real-time imaging to cells expressing different nonvisual opsins could help uncover the mechanisms underlying these phenomena (*44*). Expanding our knowledge of nonvisual opsins is important not only for understanding animal biology, but also for evaluating the impact of artificial light on wildlife (*38*, *45*, *46*).

### REFERENCES AND NOTES

1. M. Andrabi, B. A. Upton, R. A. Lang, S. Vemaraju, *Annu. Rev. Vis. Sci.* **9**, 245–267 (2023).
2. A. Terakita, *Genome Biol.* **6**, 213 (2005).
3. M. E. Guido *et al.*, *Cell. Mol. Neurobiol.* **42**, 59–83 (2022).
4. R. Karthikeyan, W. I. L. Davies, L. Gunhaga, *J. Photochem. Photobiol.* **15**, 100177 (2023).
5. Y. Nakane, T. Yoshimura, *Annu. Rev. Anim. Biosci.* **7**, 173–194 (2019).
6. K. X. Zhang *et al.*, *Nature* **585**, 420–425 (2020).
7. B. A. Gheban, I. A. Rosca, M. Crisan, *Med. Pharm. Rep.* **92**, 226–234 (2019).
8. S. Panda *et al.*, *Science* **298**, 2213–2216 (2002).
9. J. Falcón *et al.*, *Ann. N. Y. Acad. Sci.* **1163**, 101–111 (2009).
10. T. Ishikawa *et al.*, in *Laboratory Fish in Biomedical Research*, L. D'Angelo, P. de Girolamo, Eds. (Elsevier, 2022), pp. 185–213.
11. C. Umatani, M. Nakajo, D. Kayo, Y. Oka, S. Kanda, in *Laboratory Fish in Biomedical Research*, L. D'Angelo, P. de Girolamo, Eds. (Elsevier, 2022), pp. 215–243.
12. S. Kanda, in *Zebrafish, Medaka, and Other Small Fishes - New Model Animals in Biology, Medicine, and Beyond*, H. Hirata, A. Iida, Eds. (Springer Science, 2018), pp. 99–111.
13. K. Murata, M. Kinoshita, Y. Kamei, M. Tanaka, K. Naruse, Eds., *Medaka: Biology, Management, and Experimental Protocols*, vol. 2 (Wiley, 2020).
14. T.-W. Chen *et al.*, *Nature* **499**, 295–300 (2013).
15. D. O. Norris, J. A. Carr, in *Vertebrate Endocrinology* (Academic Press, ed. 6, 2021), pp. 151–204.
16. A. J. Bean, X. Zhang, T. Hökfelt, *FASEB J.* **8**, 630–638 (1994).
17. S. S. Stojilkovic, H. Zemkova, F. Van Goor, *Trends Endocrinol. Metab.* **16**, 152–159 (2005).
18. S. Martens, H. T. McMahon, *Nat. Rev. Mol. Cell Biol.* **9**, 543–556 (2008).
19. K. Sato *et al.*, *PLOS ONE* **11**, e0155339 (2016).
20. J. H. Pérez *et al.*, *Horm. Behav.* **148**, 105298 (2023).
21. T. Yamashita *et al.*, *J. Biol. Chem.* **289**, 3991–4000 (2014).
22. T. Yamashita *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22084–22089 (2010).
23. M. Jinek *et al.*, *Science* **337**, 816–821 (2012).
24. T. N. Tran, J. N. Fryer, K. Lederis, H. Vaudry, *Gen. Comp. Endocrinol.* **78**, 351–360 (1990).
25. P. Thomas, P. L. Mellon, J. Turgeon, D. W. Waring, *Endocrinology* **137**, 2979–2989 (1996).
26. K. Sato, T. Yamashita, H. Ohuchi, *J. Biol. Chem.* **299**, 105020 (2023).
27. A. Wagdi *et al.*, *Nat. Commun.* **13**, 1765 (2022).
28. K. Hosono *et al.*, *Endocrinology* **156**, 2949–2957 (2015).
29. C. Herraiz, I. Martínez-Vicente, V. Maresca, *Pigment Cell Melanoma Res.* **34**, 748–761 (2021).
30. E. M. Wolf Horrell, M. C. Boulanger, J. A. D'Orazio, *Front. Genet.* **7**, 95 (2016).
31. T. H. Nasti, L. Timares, *Photochem. Photobiol.* **91**, 188–200 (2015).
32. L. Wang *et al.*, *Aquacult. Rep.* **22**, 100946 (2022).
33. S. Fukamachi, A. Shimada, A. Shima, *Nat. Genet.* **28**, 381–385 (2001).
34. M. Spivakov *et al.*, *G3* **4**, 433–445 (2014).
35. K. Sato, K. Nwe Nwe, H. Ohuchi, *J. Comp. Neurol.* **529**, 2484–2516 (2021).
36. Y. Nakane *et al.*, *Nat. Commun.* **4**, 2108 (2013).
37. N. Nakao *et al.*, *Nature* **452**, 317–322 (2008).
38. J. D. Ni, L. S. Baik, T. C. Holmes, C. Montell, *Nature* **545**, 340–344 (2017).
39. W. L. Shen *et al.*, *Science* **331**, 1333–1336 (2011).
40. P. R. Senthilan *et al.*, *Cell* **150**, 1042–1054 (2012).
41. T. Sokabe, H.-C. Chen, J. Luo, C. Montell, *Cell Rep.* **17**, 336–344 (2016).
42. N. Y. Leung *et al.*, *Curr. Biol.* **30**, 1367–1379.e6 (2020).
43. Q. Li, N. A. DeBeaubien, T. Sokabe, C. Montell, *Curr. Biol.* **30**, 2051–2067.e5 (2020).
44. D. Whitmore, N. S. Foulkes, P. Sassone-Corsi, *Nature* **404**, 87–91 (2000).
45. B. M. Fontinha *et al.*, *PLOS Biol.* **19**, e3001012 (2021).
46. R. M. Fischer *et al.*, *PLOS Biol.* **11**, e1001585 (2013).

### ACKNOWLEDGMENTS

We thank S. Sekiguchi (University of Tokyo, Japan) for help with *pomca*:GCaMP construction. We are grateful to S. Hyodo, S. Nagata (University of Tokyo), K. Mizusawa (Kitasato University, Japan), Y. Akazome (St. Marianna University, Japan), and S. Higashijima (National Institute for Basic Biology, Japan) for helpful discussion and comments. We also thank Y. Oka (University of Tokyo) and K. Yamaguchi (RIKEN, Japan) for providing instruments. The authors thank P. Mellon (University of California–San Diego, La Jolla, CA, USA) for providing the LβT2 cell line. We are grateful to National BioResource Project (NBRP) Medaka for the gift of d-rR and Kiyosu strains. We are grateful to Vesper Studio (Tokyo, Japan) for preparation of the schematic illustrations. **Funding:** The Sasagawa Scientific Research Grant from The Japan Science Society (2020-5045 to A.F.); research grants in the Natural Sciences of Mitsubishi Foundation (to S.K.); grant for Basic Science Research Projects of Sumitomo Foundation (to S.K.); research grant from Kato Memorial Bioscience Foundation (to S.K.); research grant from the Japan Foundation for Applied Enzymology (to K.S.); and JSPS KAKENHI (grant nos. JP22J11967 to A.F., JP22KJ0833 to A.F., 18K19323 to S.K., 18H04881 to S.K., 23H02306 to S.K., 20K08885 to K.S., and 23K05850 to K.S.). **Author contributions:** Conceptualization: A.F., K.S., and S.K.; Methodology: A.F., K.S., C.F., A.T., H.O., C.U., and S.K.; Investigation: A.F., K.S., T.Y., and S.K.; Visualization: A.F., K.S., and S.K.; Funding acquisition: A.F., K.S., and S.K.; Project administration: S.K.; Supervision: S.K.; Writing – original draft: A.F., K.S., and S.K.; Writing – review and editing: A.F., K.S., T.Y., C.F., H.O., C.U., and S.K. **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.

**SUPPLEMENTARY MATERIALS**

science.org/doi/10.1126/science.adj9687
Materials and Methods
Figs. S1 to S31
Tables S1 and S2
References (*47*–*68*)
MDAR Reproducibility Checklist
Movie S1

Submitted 2 August 2023; resubmitted 10 September 2024
Accepted 1 November 2024
10.1126/science.adj9687

---

ELECTROCHEMISTRY

# Tantalum-stabilized ruthenium oxide electrocatalysts for industrial water electrolysis

**Jiahao Zhang[1,2]†, Xianbiao Fu[2]†, Soonho Kwon[3]†, Kaifeng Chen[4], Xiaozhi Liu[5], Jin Yang[6], Haoran Sun[6], Yanchang Wang[7], Tomoki Uchiyama[7], Yoshiharu Uchimoto[7], Shaofeng Li[8], Yan Li[9], Xiaolong Fan[10], Gong Chen[11], Fanjie Xia[12], Jinsong Wu[12], Yanbo Li[2], Qin Yue[2], Liang Qiao[4], Dong Su[5], Hua Zhou[9], William A. Goddard III[3]*, Yijin Kang[1]***

**The iridium oxide ($IrO_2$) catalyst for the oxygen evolution reaction used industrially (in proton exchange membrane water electrolyzers) is scarce and costly. Although ruthenium oxide ($RuO_2$) is a promising alternative, its poor stability has hindered practical application. We used well-defined extended surface models to identify that $RuO_2$ undergoes structure-dependent corrosion that causes Ru dissolution. Tantalum (Ta) doping effectively stabilized $RuO_2$ against such corrosion and enhanced the intrinsic activity of $RuO_2$. In an industrial demonstration, Ta-$RuO_2$ electrocatalyst exhibited stability near that of $IrO_2$ and had a performance decay rate of ~14 microvolts per hour in a 2800-hour test. At current densities of 1 ampere per square centimeter, it had an overpotential 330 millivolts less than that of $IrO_2$.**

Water electrolysis driven by renewable power sources is an important route for green hydrogen production (*1*). Unfortunately, the sluggish kinetics, low efficiency, and excess energy consumption required for oxygen evolution reaction (OER) make OER the bottleneck for high-efficiency water electrolysis (*2*, *3*). Proton-exchange membrane water electrolysis (PEM-WE) is the most attractive among all water electrolysis technologies for its high current density (and hence high hydrogen production rate), high purity, and high pressure for hydrogen output and quick response that enables ready coupling with renewable power sources (*4*, *5*). However, the harsh operating environment (high acidity and high potential) of OER electrocatalysts poses severe requirements in chemical stability for PEM-WE (*6*–*8*). Although many new OER electrocatalysts have been discovered (*8*–*20*), the PEM-WE industry still primarily uses first-generation iridium oxide ($IrO_2$), which is limited for large-scale applications by the extreme scarcity and high cost of Ir (*21*, *22*). Additionally, although the stability of $IrO_2$ meets current industrial requirements, the activity of $IrO_2$ OER electrocatalysts must be dramatically improved to enable competitive hydrogen production.

Ruthenium oxide ($RuO_2$), the second-most attractive option for OER catalysts, offers advantages, including the relatively high reserves and low cost of Ru compared with Ir, as well as high activity (*23*–*26*). However, the poor stability of $RuO_2$ prevents practical applications. The Faradaic efficiency of Ru dissolution for $RuO_2$ electrocatalysts in the normal operation of water electrolysis is ~0.1%, which is at least two orders of magnitude higher than that of the Ir dissolution in $IrO_2$ catalysts. To improve the performance of $RuO_2$, it is crucial to identify new modified materials that extend the structure-property correlation for both activity and stability, as well as to establish the reaction mechanism. For example, many studies suggest that lattice oxygen exchange is involved in the OER process and is responsible for the dissolution of $RuO_2$ catalysts (*17*, *18*, *27*), whereas others claim lattice oxygen exchange is not observed (*28*). Mixing tantalum pentoxide ($Ta_2O_5$) and $RuO_2$ (or $IrO_2$) has led to slightly improved stability (*29*, *30*), but this approach has not been further explored.

In this study, we examined well-defined extended surfaces of $RuO_2$ with selected crystalline orientations to investigate the structure-property correlations in OER. We identified structure-dependent corrosion of $RuO_2$ as responsible for both low stability and low activity. However, we found that doping Ta into $RuO_2$ both enhanced the intrinsic activity of $RuO_2$ toward OER and suppressed $RuO_2$ corrosion, while also alleviating Ru dissolution in water electrolysis, dramatically improving electrocatalyst stability. To understand the atomistic origin of how Ta improved performance and stability, we performed grand canonical quantum mechanics (GCQM) calculations (*31*) to describe the mechanism and stability as a function of applied potential. Based on what we learned from these QM and experimental studies on the well-defined extended surface of $RuO_2$, we synthesized nanoparticulate Ta-$RuO_2$ electrocatalysts and demonstrated performance in an industrial-scale membrane electrode assembly (MEA). Moreover, a full industrial demonstration showcased hydrogen production from solar energy at up to megawatt rates to feed a hydrogen refueling station.

## Well-defined extended surface of $RuO_2$

$RuO_2$ films were prepared by pulse laser deposition (demonstrated in supplementary materials, materials and methods). The films used in our experiments showed similar thickness and root mean square (RMS) roughness according to x-ray reflection (XRR) (fig. S1) and atomic force microscopy (fig. S2 and table S1). Before the OER experiments, we examined the cathodic pseudocapacitive charge, $q^*_{cathodic}$ (*32*, *33*), of the thin films to determine the population of available electrochemically active sites (fig. S3A and table S2). Because $RuO_2$ extended surfaces with different orientations have different surface atomic arrangements (Fig. 1A, insets), the number of coordinatively unsaturated sites (CUS) of Ru for various surfaces is expected to be different. We calculated the CUS as 10 $nm^{-2}$ for (001), 7 $nm^{-2}$ for (100), 5 $nm^{-2}$ for (110), and 7 $nm^{-2}$ for (111). As shown in Fig. 1B, the measured cathodic charge densities showed a good linear correlation with a Pearson correlation coefficient of 0.91 to the

---

[1]School of Chemical Engineering, Sichuan University, Chengdu, China. [2]Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China. [3]Materials and Process Simulation Center, California Institute of Technology, Pasadena, CA, USA. [4]School of Physics, University of Electronic Science and Technology of China, Chengdu, China. [5]Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing, China. [6]DongFang Boiler Group Co., LTD, Chengdu, China. [7]Graduate School of Human and Environmental Studies, Kyoto University, Sakyo-ku, Kyoto, Japan. [8]Department of Physics, Technical University of Denmark, Kongens Lyngby, Denmark. [9]X-ray Science Division, Advanced Photon Source, Argonne National Laboratory, Lemont, IL, USA. [10]The Key Laboratory for Magnetism and Magnetic Materials, Lanzhou University, Lanzhou, China. [11]National Laboratory of Solid State Microstructures, Department of Physics and Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing, China. [12]State Key Laboratory of Advanced Technology for Materials Synthesis and Processing, School of Materials Science and Engineering, Wuhan University of Technology, Wuhan, China.
***Corresponding author. Email: kangyijin@scu.edu.cn (Y.K.); wag@caltech.edu (W.A.G.)**
†These authors contributed equally to this work.

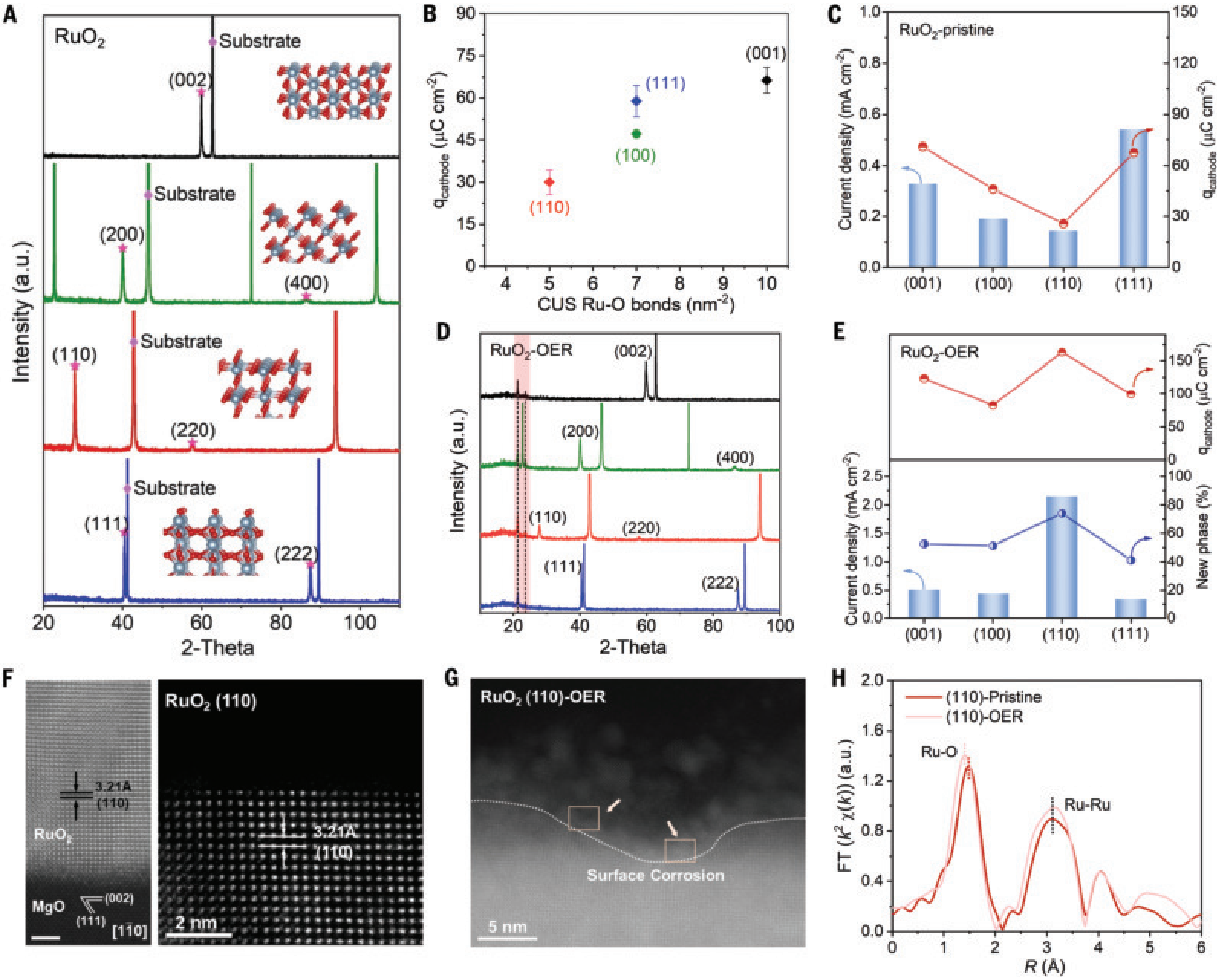


**Fig. 1. Investigation on well-defined extended surfaces of $RuO_2$.** (**A**) XRD patterns of thin films with selected orientations that expose $RuO_2$(001), $RuO_2$ (100), $RuO_2$ (110), and $RuO_2$ (111) surfaces, respectively. Red stars highlight the characteristic peaks for each orientation, and purple rhombuses highlight the peaks for substrate materials. Insets in (A) illustrate the structure for selected crystalline orientations. (**B**) The $q^*_{cathodic}$ for each orientation as a function or CUS Ru-O number. (**C**) Bar chart showing apparent current densities and the $q^*_{cathodic}$ for each surface. (**D**) XRD patterns for thin films after electrochemical activation, showing that a new phase evolves, as highlighted in the shaded pink region. (**E**) Chart showing the correlation among activity (current density), $q^*_{cathodic}$, and the phase ratio for the thin films after activation. All current density bar charts are at 1.6 V versus a reversible hydrogen electrode (RHE). (**F**) Cross-sectional HAADF-STEM images of $RuO_2$(110) film, indicating that $RuO_2$ film grew epitaxially along the surface of MgO and terminated by a near atomically flat surface. Scale bar, 2 nm. (**G**) HAADF-STEM image of $RuO_2$(110) surface after OER test. (**H**) Fourier-transform (FT) extended x-ray absorption fine structure of $RuO_2$(110), obtained by grazing incidence x-ray absorption spectrum.

CUS Ru numbers [which is consistent with previous reports (*28*)], indicating that the surface-exposed Ru atoms were available for electrochemical reactions.

The OER properties of $RuO_2$ thin films were evaluated first by cyclic voltammetry (CV), as shown in fig. S4. The OER current densities of pristine $RuO_2$ surface correlated strongly with the charge density, $q^*_{cathodic}$, indicating that the population of Ru sites was a dominant factor for the overall OER activity of $RuO_2$ thin films (Fig. 1C). We note the electrochemical properties of $RuO_2$ thin films that changed after CV cycling: Although (001), (110), and (100) surfaces showed an apparent activation process, the (111) surface exhibited a decay in current density over time but no activation barrier (fig. S4).

We used x-ray diffraction (XRD) to examine the thin films after 10 cycles of CV to understand the changes in electrode performance (Fig. 1D). Two new XRD peaks at 21.2° and 23.6° appeared on all thin films regardless of orientation. The best match for the new peaks is $RuO_4$ [monoclinic, C2/c (*15*)], which has been suggested to be involved in the corrosion process of $RuO_2$. (*34*, *35*). We calculated the ratio of the new $RuO_4$ phase for each orientation in the XRD patterns to determine the extent of $RuO_2$ thin film evolution during OER. As shown in Fig. 1E, the current densities of $RuO_2$ after the activation process for different crystalline orientations correlated strongly with both the phase ratio and $q^*_{cathodic}$ (i.e., electrochemical surface area). The trend of $q^*_{cathodic}$ values to crystalline orientations was completely different from that of thin films before OER activation (compare Fig. 1C and Fig. 1E), indicating that the change in surface area was structure dependent.

We used aberration-corrected transmission electron microscope and synchrotron light source–based techniques to track the structure evolution of $RuO_2$ thin films during OER. Because $RuO_2(110)$ exhibited the most drastic change in OER reaction among all the selected orientations, we chose it for the demonstration. Energy-dispersive spectroscopy (EDS) (fig. S5A) maps and high-angle annular dark-field scanning transmission election microscopy (HAADF-STEM) (fig. S5B) of the cross-section of pristine $RuO_2(110)$ film presented a highly crystalline and long-range ordered (110) plane with 47 nm in thickness, which was consistent with results from XRR and XRD. Atomic-resolution HAADF-STEM (as shown in Fig. 1G) revealed that $RuO_2(110)$ epitaxially grew on MgO (001) and terminated by an atomically flat surface.

Low-energy electron microscopy (fig. S6) revealed a uniform film morphology for $RuO_2$ (110), confirming the atomically flat surface. However, the film underwent severe corrosion during OER. Synchrotron XRR data-fitting results indicated that the roughness of $RuO_2(110)$ surface changes from 1.3 nm for the pristine film to 5.5 nm after OER (fig. S7, A and B, and table S3) with a drastic increase of $q^*_{cathodic}$, suggesting a substantially increased population of active sites and surface reconstruction. Scattering-length density analysis (fig. S7C) demonstrated that an uneven and lower mass density layer formed on the (110) film after OER, which correlated with Ru leaching and severe surface reconstruction. XRD confirmed the existence of a new phase (fig. S8A).

We used grazing incidence diffraction (GID) as a surface-sensitive probe to identify the new phase as a species protruding out of the surface plane, most likely as high-valent Ru oxide clusters (fig. S8B). The broad peak width in XRD and GID suggested that this new structural phase had low crystallinity. Additionally, the substantially reduced intensities for $RuO_2$ (110), $RuO_2(220)$, and MgO(004) in the GID pattern also imply the high epitaxial quality of the thin film samples (e.g., Bragg conditions for substrate and epitaxial film were not satisfied at the GID scattering geometry). Figure 1G and fig. S9 show the eroded surface structure of $RuO_2(110)$ after OER, where surface corrosion accompanied with formation of polycrystalline and amorphous $RuO_x$ species was observed.

We also found Ru sites transforming to disordered phases at many step sites (as shown in the pink box in fig. S9), implying that low-coordinated step sites might be the trigger points of continuous degradation. The surface species were further revealed with the help of grazing-incidence x-ray absorption spectrum (GIXAS). In the x-ray absorption near-edge structure spectra (fig. S10), the thin film after OER had a higher edge position and greater white-line intensity than the pristine film, which implies increased valence for Ru after OER. Fourier transform of the extended x-ray absorption fine structure revealed a slightly reduced Ru-O bond length for the film after OER (Fig. 1H), which corresponded to a shorter Ru-O bond (1.7 Å) of $RuO_4$ rather than that of $RuO_2$ (1.94 or 1.98 Å).

From the above experimental results, we concluded that under OER conditions, a fast and structure-dependent corrosion process started at the surface of $RuO_2$ electrocatalysts to produce high-valent Ru species and that this corrosion process activated $RuO_2$ toward OER through the substantially increased electrochemical surface area of the corroded surface. Although this corrosion improved the OER activity of $RuO_2$ initially through an apparent activation, the high Ru dissolution rate from corrosion leads to failure of the electrocatalyst in a relatively short period.

## Ta-modified $RuO_2$

For practical application in the PEM-WE industry, the lifetime of the $RuO_2$ catalyst must be dramatically improved before $RuO_2$ can be used as an OER electrocatalyst. We explored transition metal–modified $RuO_2$ materials for improved activity, stability, or both. We identified that Ta improved both the intrinsic activity and the chemical stability of $RuO_2$ catalysts.

We added Ta onto $RuO_2$ thin films through magnetron sputtering. Adding Ta atoms on the well-defined $RuO_2$ surface did not change the crystalline orientation, as shown by XRD (Fig. 2A and fig. S11). A slight increase in RMS roughness was observed on Ta-doped films through atomic force microscopy (fig. S12 and table S4). The cross-sectional HAADF-STEM image of $TaRuO_2(110)$ film in Fig. 2B depicts a well-defined interface between the film and MgO(001) substrate and also a well-preserved and ordered (110) plane after Ta sputtering, confirming that Ta sputtering did not damage the film orientation.

The HAADF-STEM image in Fig. 2B also shows an ordered and near atomically flat surface lattice of $TaRuO_2(110)$ film after Ta sputtering. The distance between the atoms highlighted with red circles and the subsurface Ru atoms (3.39 to ~3.40 Å) is larger than that of the adjacent atoms and the subsurface atoms (3.21 Å), which is likely caused by Ta atom occupation. The electron energy loss spectra (EELS) of surface atomic layers (Fig. 2C and fig. S13), marked with an orange box in Fig. 2B, showed a distinct signal at 1780 to ~1795 eV, which corresponds to the Ta-$M_4$ edge, whereas no EELS Ta signal was observed in the $RuO_2$ (110). The evidence above indicated that Ta atoms were successfully doped in the surface $RuO_2$ lattice after sputtering, whereas the surface orientation and flatness were nearly unchanged. X-ray photoelectron spectroscopy (fig. S14) of these Ta-doped films showed that the binding energy of Ta 4f is ~26.3 eV, revealing a +5 valence state of Ta on $RuO_2$ films. The amount of Ta doping is also detailed in table S5.

After the $RuO_2$ surface was modified by Ta atoms, the OER activity improved dramatically. Figure 2D and fig. S15 show the CV curves with iR correction for Ta-doped and pure $RuO_2$ films. Each Ta-doped film showed much larger current densities of OER compared with corresponding pure $RuO_2$ film. For example, Ta-$RuO_2$ (110) exhibited a geometric current density of 0.61 mA $cm^{-2}$ at 1.6 V versus reference hydrogen electrode (RHE), which was five-times higher than that of $RuO_2$ (110). Tafel plots in fig. S16 demonstrate that Ta-doped $RuO_2$ films showed lower Tafel slopes compared with pure $RuO_2$, namely 78.3 mV $dec^{-1}$ (where "dec" is "decade") for Ta-$RuO_2$ (110) and 131.2 mV $dec^{-1}$ for $RuO_2$ (110), indicating that Ta-doped $RuO_2$ films had the faster kinetic rate of OER. We also studied the electrochemical impedance spectroscopy (EIS) to compare the rate of OER between Ta-doped and pure $RuO_2$ films. The Nyquist plots for Ta-doped and pure $RuO_2$ films at 1.6 V versus RHE are presented in fig. S17. According to the fitting results (table S6), Ta-doped $RuO_2$ films exhibited much smaller charge transfer resistance ($R_{ct}$) compared with pure $RuO_2$ films, i.e., $R_{ct}$ = 478 ohms for Ta-$RuO_2$ (110) and $R_{ct}$ = 3392 ohms for $RuO_2$ (110), indicating that Ta-doped $RuO_2$ films provide faster charge transfer and rate of reaction toward OER. The change in capacitance charge over CV cycles under OER conditions was also greatly suppressed (fig. S18). Although the activities of pristine Ta-$RuO_2$ thin films with various crystalline orientations still correlated with the pseudocapacitive charge densities (fig. S19), this correlation was weaker than that for $RuO_2$ thin films, which implied that Ta changed the intrinsic activity of $RuO_2$, in addition to suppressing the morphology and structure evolution of electrocatalysts (fig. S20).

As shown in Fig. 2E, the surface of $TaRuO_2$ (110) thin films was less drastically changed after the OER test than that of $RuO_2(110)$. EELS spectra of $TaRuO_2(110)$ surface region (fig. S21) further verified that Ta remained in the surface lattice after the OER test, confirming the stronger structure stability of the Ta-doped $RuO_2$ surface. Accordingly, less new phase formation (high-valent Ru oxide species) was observed on Ta-modified $RuO_2$ thin films after OER (Fig. 2F). In situ and ex situ electrochemical Raman spectroscopy analysis (fig. S22) offered more evidence about degradation of $RuO_2$ and verified that Ta doping substantially improved the stability of $RuO_2$ films. We clearly observed the vibration modes of $E_g$ (536 $cm^{-1}$), $A_{1g}$ (649 $cm^{-1}$), and $B_{2g}$ (715 $cm^{-1}$), which are ascribed to the featured modes of $RuO_2$ (110) under open circuit potential (OCP) in fig. S22A. Two distinct vibration modes also appeared around 428 $cm^{-1}$

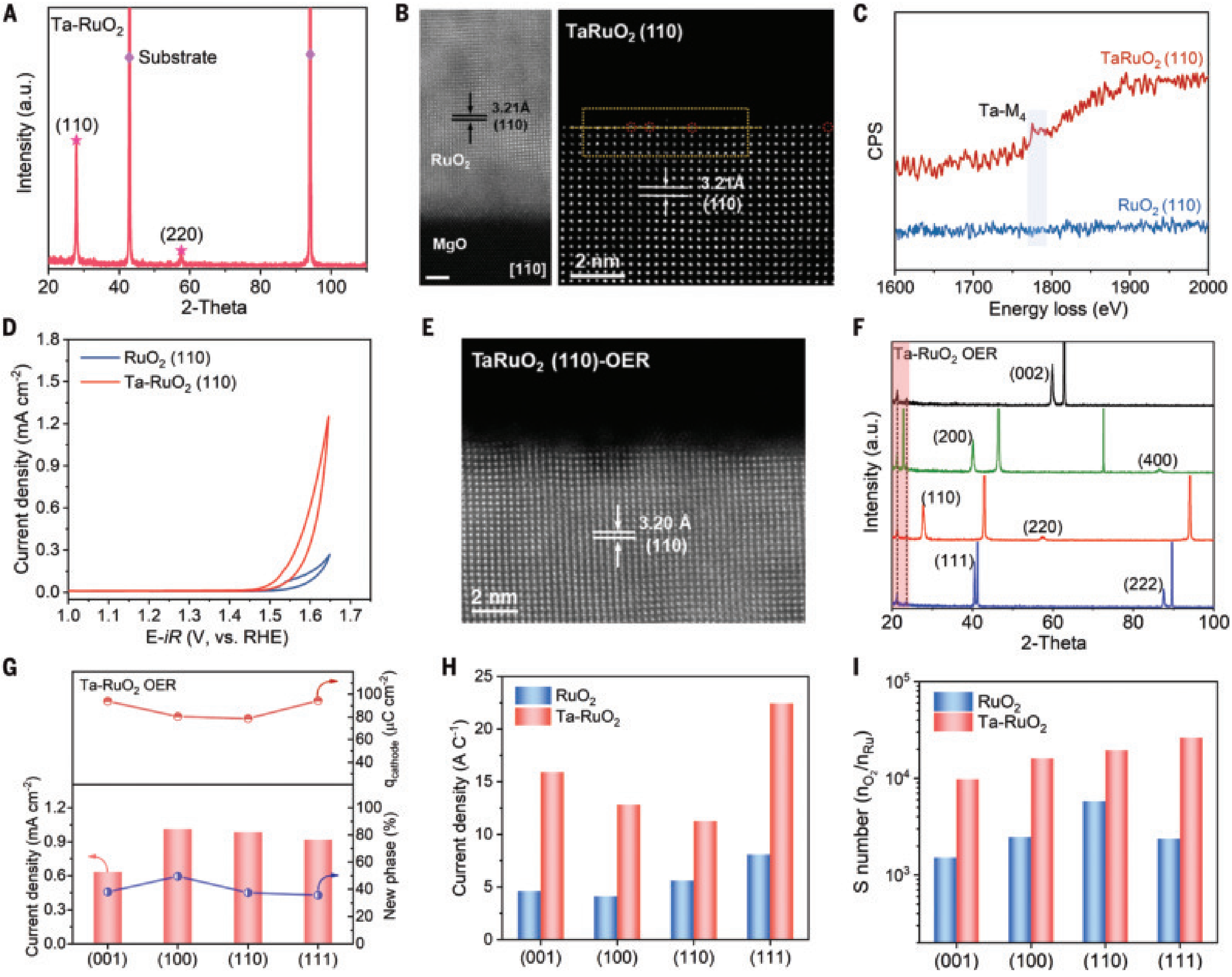


**Fig. 2. Investigations on Ta-modified $RuO_2$ surfaces.** (**A**) XRD pattern of Ta-doped $RuO_2$(110) thin film. Red stars highlight the characteristic peaks for $RuO_2$(110), and purple rhombuses highlight the peaks for the substrate material. (**B**) Cross-sectional HAADF-STEM images of Ta-doped $RuO_2$(110) film, indicating that Ta sputtering does not damage the film orientation or crystallization. Scale bar, 2 nm. (**C**) EELS spectra of Ta-$M_4$ edge. CPS, counts per second. (**D**) CV curves of $RuO_2$ (110) and Ta-$RuO_2$ (110). (**E**) HAADF-STEM image of $TaRuO_2$(110) surface after OER test. (**F**) XRD patterns for thin films after electrochemical activation, showing that a new phase evolves, as highlighted. (**G**) Chart showing the correlation among activity (current density), $q^*_{cathodic}$, and the phase ratio for the thin films after activation. (**H**) Bar chart showing current densities normalized by $q^*_{cathodic}$ to represent intrinsic activities. (**I**) Bar chart showing the thin film *S* number, which is defined as the molar ratio of $O_2$ produced to dissolved Ru, representing the stability of electrocatalysts. All bar charts show current density at 1.6 V versus RHE.

and 594 $cm^{-1}$, which can be ascribed to the bond between Ru sites and hydroxyl intermediates (*36*). With increasing applied potentials, the vibration modes of $E_g$, $A_{1\,g}$, and $B_{2\,g}$ exhibited apparent red shifts toward low wave number, while the vibration modes of Ru-OH remained unchanged. For example, the $E_g$ mode showed a red shift from 536 $cm^{-1}$ to 524 $cm^{-1}$ when the applied potential reached 1.73 V versus RHE. The red shifts in peak position of $RuO_2$ (110) can be attributed to the formation of nanosized clusters during OER (*37*), implying the reconstruction of $RuO_2$ (110) single-crystal film, which is consistent with our high-resolution transmission electron microscopy (HRTEM) results. Apart from the red shifts of vibration modes, the intensity of the $E_g$, $A_{1\,g}$, and $B_{2\,g}$ modes of $RuO_2$ (110) dramatically decreased with increasing applied potentials. Specifically, the $A_{1\,g}$ and $B_{2\,g}$ modes of $RuO_2$ (110) almost disappeared when the potential reached 1.73 V versus RHE, showing the destruction of Ru-O configurations in $RuO_2$ (110) crystal structure during OER. The in situ Raman spectra of Ta-$RuO_2$ is illustrated in fig. S22B (110). The $E_g$, $A_{1\,g}$, and $B_{2\,g}$ modes of Ta-$RuO_2$ (110) are located at 537 $cm^{-1}$, 649 $cm^{-1}$, and 716 $cm^{-1}$ under OCP, respectively, which are similar to that of pure $RuO_2$ (110). The vibration of Ru-OH also appeared around 429 $cm^{-1}$ and 594 $cm^{-1}$. Compared with the apparent changes in Raman spectra of $RuO_2$ (110) during OER, Ta-$RuO_2$ (110) showed few variations in Raman signals with increasing applied potentials. The $E_g$ mode of Ta-$RuO_2$ (110) stayed at 537 $cm^{-1}$ under all the applied potentials in our experiments, which shows that doping Ta noticeably prevents the $RuO_2$ (110) surface from reconstruction and suppresses the destruction of $RuO_2$ crystal structure. Considering the strong fluorescence below 300 $cm^{-1}$ under in situ conditions, we carried out ex situ Raman characterizations (as shown in fig. S22, C and D). We found similar red shifts for the $E_g$, $A_{1\,g}$, and $B_{2\,g}$ modes of $RuO_2$ (110) after OER (fig. S22C). Moreover, a distinct Raman peak appeared around 229 $cm^{-1}$, which is ascribed to $RuO_4$ configurations (*38*) on $RuO_2$ (110) film after OER (fig. S22C). This observation implies the formation of high-valent Ru species on $RuO_2$ films, which is consistent with the XRD and GIXAS results in Fig. 1. However, we observed neither the red shifts of vibration modes nor the vibrations of $RuO_4$ on Ta-$RuO_2$ (110) after OER (fig. S22D). Although the activities of Ta-$RuO_2$ still correlated with both $RuO_4$ phase ratio and capacitance charge (Fig. 2G), the correlation for Ta-$RuO_2$ no longer followed the same trend seen in $RuO_2$ (fig. S23). After normalizing the current densities by the corresponding $q^*_{cathodic}$ to reflect the intrinsic OER activities (Fig. 2H), Ta doping was shown to improve the intrinsic activity of $RuO_2$ toward OER.

The stability of $RuO_2$ and Ta-$RuO_2$ was quantified by measuring Ru dissolution under continuous operation of OER at 1.6 V, as determined by inductively coupled plasma mass spectrometry (ICP-MS). The Ru dissolution rates for respective crystal faces were 11.9 (001), 5.3 (100), 44.5 (110), and 5.7 (111) ng $hour^{-1}$ for $RuO_2$, and 3.9 (001), 2.4 (100), 21.8 (110), and 2.4 (111) ng $hour^{-1}$ for Ta-$RuO_2$, respectively (figs. S24 and S25). This trend of Ru dissolution in $RuO_2$ was consistent with previous reports. (*39*) However, the Ta modification suppressed Ru dissolution in $RuO_2$ electrocatalysts. The *S* number, (*40*) which is defined as the molar ratio of $O_2$ produced to dissolved Ru, is calculated and presented in Fig. 2I, illustrating the much-improved activity and stability of $RuO_2$ electrocatalyst resulting from Ta doping.

### Computational studies

We performed a series of QM calculations to provide insight into the atomistic origin of the improved intrinsic activity and stability of $RuO_2$ from Ta incorporation. We found that the (110) surface led to the most drastic change during the apparent activation, whereas (100) led to the least drastic change. Here we used the GCQM method (*31*) to compare the activity of (110) and (100) surfaces as a function of applied potential with and without surface Ta doping. The atomic surface geometries and distinctive surface sites for $RuO_2$ (100) and (110) facets are shown in figs. S26, A and B.

Under electrochemical conditions, applying anodic polarization led to the successive oxidation of the surface state. Starting from the surface fully covered by hydroxyl groups, we constructed the surface free-energy diagrams for pure $RuO_2$ (fig. S26, C and D). The surfaces are saturated by bridge oxo ($\mu_2$-O) and by terminal oxo ($\mu_1$-O) at the potential of 1.31 V for $RuO_2$(100) and at the potential of 1.21 V versus RHE for $RuO_2$(110), implying the facile surface oxidation on the $RuO_2$ surface. For Ta-$RuO_2$, we found that Ta preferentially doped at the surface on $RuO_2$(100) (fig. S27), where only CUS sites were present. The Ta-$RuO_2$ (100) surface showed a dominant Ta-OH state even at a high potential (fig. S26E). However, although Ta preferred the surface bridge site (BRI) over the CUS site (fig. S28), we focused on the role of Ta at the CUS because the Ta at the BRI site was not only directly involved in the OER pathway but also decreased OER activity (fig. S29). For Ta-$RuO_2$(110) with $Ta_{CUS}$, the hydroxyl on Ta ($\mu_1^{Ta}$-OH) was slightly more favorable than the terminal oxo ($\mu_1^{Ta}$-O) at potentials up to 1.65 V versus RHE, indicating a higher population of $\mu_1^{Ta}$-OH (fig. S26F).

The reaction mechanism for OER on $RuO_2$ (100) is shown in Fig. 3A. A surface $H_2O$ molecule reacted with the $\mu_1$-O to make a new O–O bond forming $\mu_1$-OOH (state 2) and simultaneously transferred its H to an adjacent $\mu_2$-O to form $\mu_2$-OH. This reaction had a reaction free energy ($\Delta G$) of 0.40 eV and a free-energy barrier, $\Delta G^{\ddagger}$, of 0.77 eV at 1.6 V versus RHE (Fig. 3B). After making $\mu_1$-OOH, the successive exothermic deprotonation and $O_2/H_2O$ exchange steps returned the surface back to state 1. On Ta-$RuO_2$(100), the proton was transferred to $\mu_1^{Ta}$-OH rather than $\mu_2$-O during O–O bond formation, which led to formation of $\mu_1^{Ta}$-$H_2O$ and $\mu_1$-OOH on Ru (Fig. 3A, bottom). This reaction had a 0.07-eV-lower $\Delta G^{\ddagger}$ of 0.70 eV at 1.6 V versus RHE compared with $RuO_2$(100) (Fig. 3B).

On the $RuO_2$(110) surface, although the $\mu_2$-O was a stronger proton acceptor (fig. S26D and S30), the kinetic barrier of the water dissociation step transferring the proton to $\mu_1$-O had a 0.09-eV-lower barrier of 0.78 eV at 1.6 V versus RHE. The Ta doping at CUS leads to much more favorable water dissociation with $\Delta G^{\ddagger}$ = 0.57 eV and $\Delta G$ = 0.30 eV at 1.6 V versus RHE (Fig. 3C).

To predict the current density, we calculated the turnover frequency (TOF) by assuming that the O–O coupling step is the rate-determining step (supplementary text). We predicted that $RuO_2$(100) had a 1.8-fold higher TOF of 0.69 $s^{-1}$ compared with $RuO_2$(110), with a TOF of 0.36 $s^{-1}$ (Fig. 3D). This analysis led to a current density of 0.32 mA $cm^{-2}$ for $RuO_2$(100) and 0.12 mA $cm^{-2}$ for $RuO_2$(100) (fig. S31), which was consistent with our experiments (Fig. 1C and fig. S4E). The Ta doping on $RuO_2$(100) led to a TOF of 5.62 $s^{-1}$ at 1.6 V versus RHE, which was 8.1-fold faster than pure $RuO_2$(100). This analysis led to a current density of 0.69 mA $cm^{-2}$, which was more than two times higher than that of pure $RuO_2$(100), for a doping concentration of 8.9%, which again agreed very well with our experiment in which the Ta-doped $RuO_2$ (100) had a current density two times higher than that of pure $RuO_2$ (Fig. 2H).

We predicted that Ta doping on $RuO_2$(110) could have dramatic effects. The TOF near the $Ta_{CUS}$ was calculated as 382 $s^{-1}$ at 1.6 V versus RHE, which is 1000 times faster than for pure $RuO_2$(110). However, this $Ta_{CUS}$ was thermodynamically less stable than $Ta_{BRI}$. Therefore, there may be only trace amounts of this phase present. We estimate that the presence of only 0.04 atomic % (at %) of surface $Ta_{CUS}$ could lead to a current density 1.8 times higher than that of pure $RuO_2$(110), as observed in our experiment (Fig. 2H).

To provide a theoretical explanation for the enhanced durability of Ta-$RuO_2$ at OER working conditions, we calculated the energetics of the dissolution process. Based on our experiment, we conclude that the corrosion most likely involves the high-valent Ru clusters (figs. S8 to S10). Using ab initio molecular dynamics, Klyukin *et al.* suggested that Ru dissolution occurs through deprotonation of the $RuO_2(OH)_2$ cluster (*34*). Because corrosion occurs mainly at the defect site (*41*), we built a high-index (231) surface model of rutile structure (fig. S32A) that exposed the (110) surface with the $[1\bar{1}1]$ step edge having the alternative two metal sites, $M_{BRI}$ and $M_{CUS}$ (fig. S32B). There were three distinctive $\mu_1$-O motifs to bind the cluster on the edge. One motif was bonded to $M_{BRI}$ ($M_{BRI}$-O), and the other two motifs were bonded to $M_{CUS}$, which were either parallel ($M_{CUS}$-$O^{=}$) or perpendicular ($M_{CUS}$-$O^{\perp}$) to the (110) terrace plane.

We calculated the energetics of the $RuO_2(OH)_2$ oxidation on these different binding sites, which led to Ru dissolution (*34*). Figure 3E (top) shows the representative oxidation process on the $Ru_{CUS}$-$O^{\perp}$ site on pure $RuO_2$, and the energies are shown in Fig. 3F. The reaction energies of the deprotonation of $RuO_2(OH)_2$ complex ($\Delta G^{1st}$) were 0.97 eV, 1.04 eV, and 1.28 eV for the $Ru_{BRI}$-O, $Ru_{CUS}$-$O^{=}$, and $Ru_{CUS}$-$O^{\perp}$ sites, respectively (table S7). The deprotonation led to the bond-length increase between Ru in the cluster and $\mu_1$-O ($d_{Ru\text{-}O}$), leading to the $RuO_3OH$ detachment except for the bond at $Ru_{BRI}$-O sites, which still maintained a chemical bonding with $d_{Ru\text{-}O}$ of 2.00 Å. However, the following second deprotonation led to Ru dissolution without exception (table S8), with much lower energy ($\Delta G^{2nd}$), which is in agreement with the previous study (*34*).

To investigate the role of Ta dopant on durability, we built corresponding (231) surfaces with different doping sites, $Ta_{CUS}$ (fig. S32C) and $Ta_{BRI}$ (fig. S32D). Although the Ta dopant still preferred the CUS site over BRI sites on the edge by −0.81 eV on the intact edge structure, we noted that the $Ta_{CUS}$ was stabilized by the binding of the $RuO_2(OH)_2$ cluster. This strong interaction left the cluster with even lower energy by 0.39 eV for $Ta_{CUS}$-$O^{\perp}$ and 0.11 eV for $Ta_{CUS}$-$O^{=}$ sites compared with the $Ta_{BRI}$ counterparts. The $\Delta G_1$ for the deprotonation of the intermediate were 1.25 eV for $Ta_{CUS}$-$O^{\perp}$ and 1.26 eV $Ta_{CUS}$-$O^{=}$, which were higher by 0.28 eV and 0.22 eV compared with the corresponding cases on pure $RuO_2$(231) facets.

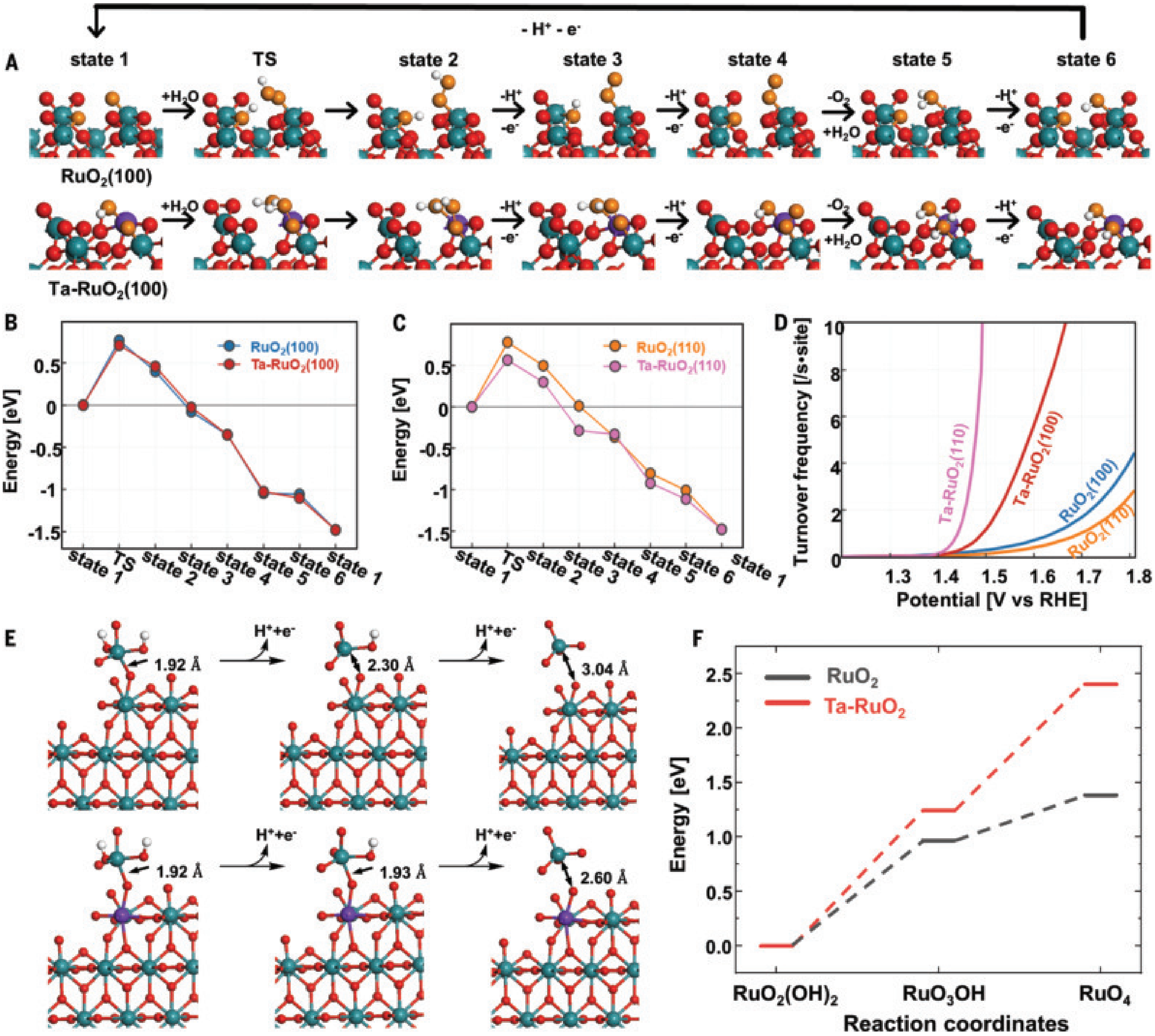


**Fig. 3. OER reaction mechanisms over $RuO_2$ and Ta-doped $RuO_2$.** (**A**) The OER pathway on pure $RuO_2(100)$ and Ta-doped $RuO_2(100)$. TS denotes the transition state. The proton is transferred to a terminal oxo during the water dissociation step except for the pure $RuO_2(100)$. The purple, turquoise, red, orange, and white spheres denote Ta, Ru, O, and O involved in a single OER cycle and H atoms, respectively. The free-energy landscapes along the OER reaction coordinates on the (**B**) (100) surface and (**C**) (110) surface. The blue, red, yellow, and pink colors represent the pure $RuO_2$ (100), Ta-doped $RuO_2(100)$, pure $RuO_2$ (110), and Ta-doped $RuO_2$ (110), respectively. (**D**) Turnover frequency (TOF) on the four model surfaces. (**E**) Ru dissolution on the step edge of the $RuO_2$ and Ta-doped $RuO_2$ (231) surface. The serial oxidation of Ru dissolution intermediate, $RuO_2(OH)_2$, bound to the edge CUS site through $\mu_1$-$O^{\perp}$ is depicted on pure $RuO_2$ (top) and on Ta-$RuO_2$ (bottom). (**F**) The potential energy landscape during the two sequential deprotonations, which shows that the Ru dissolution on Ta-$RuO_2$ requires more energy compared with pure $RuO_2$.

Moreover, the oxidized intermediate, $RuO_3OH$, still formed a chemical bond to the $\mu_1$-O with a $d_{\text{Ru-O}}$ of 1.93 Å for $Ta_{CUS}$-$O^{\perp}$ and 1.83 Å for $Ta_{CUS}$-$O^{=}$. The $\Delta G_2$ values toward $RuO_4$ dissolution of 1.16 eV and 1.05 eV, respectively, were 0.74 and 0.77 eV higher than that on pure $RuO_2$. We could not find such a stability increase of the bound intermediate with Ta doping on the BRI site (table S7), suggesting that the strong binding of the Ru dissolution intermediate on $Ta_{CUS}$-O enhanced the durability of the Ta-$RuO_2$ catalyst in the working condition.

## Demonstration of Ta-$RuO_2$ for industrial water electrolysis

Based on what we learned from studying the above well-defined extended surfaces, we implemented the synthesis of nanoparticulate Ta-$RuO_2$ electrocatalysts and applied them in an industrial environment to demonstrate Ta-$RuO_2$ as a practical electrocatalyst for industrial water electrolysis. Such Ta-$RuO_2$ electrocatalysts have been commercialized, mass produced, and are available on the market. We synthesized the nanoparticulate Ta-$RuO_2$ materials through a modified sol-gel method. (*24*, *42*) A mixture of ruthenium chloride and

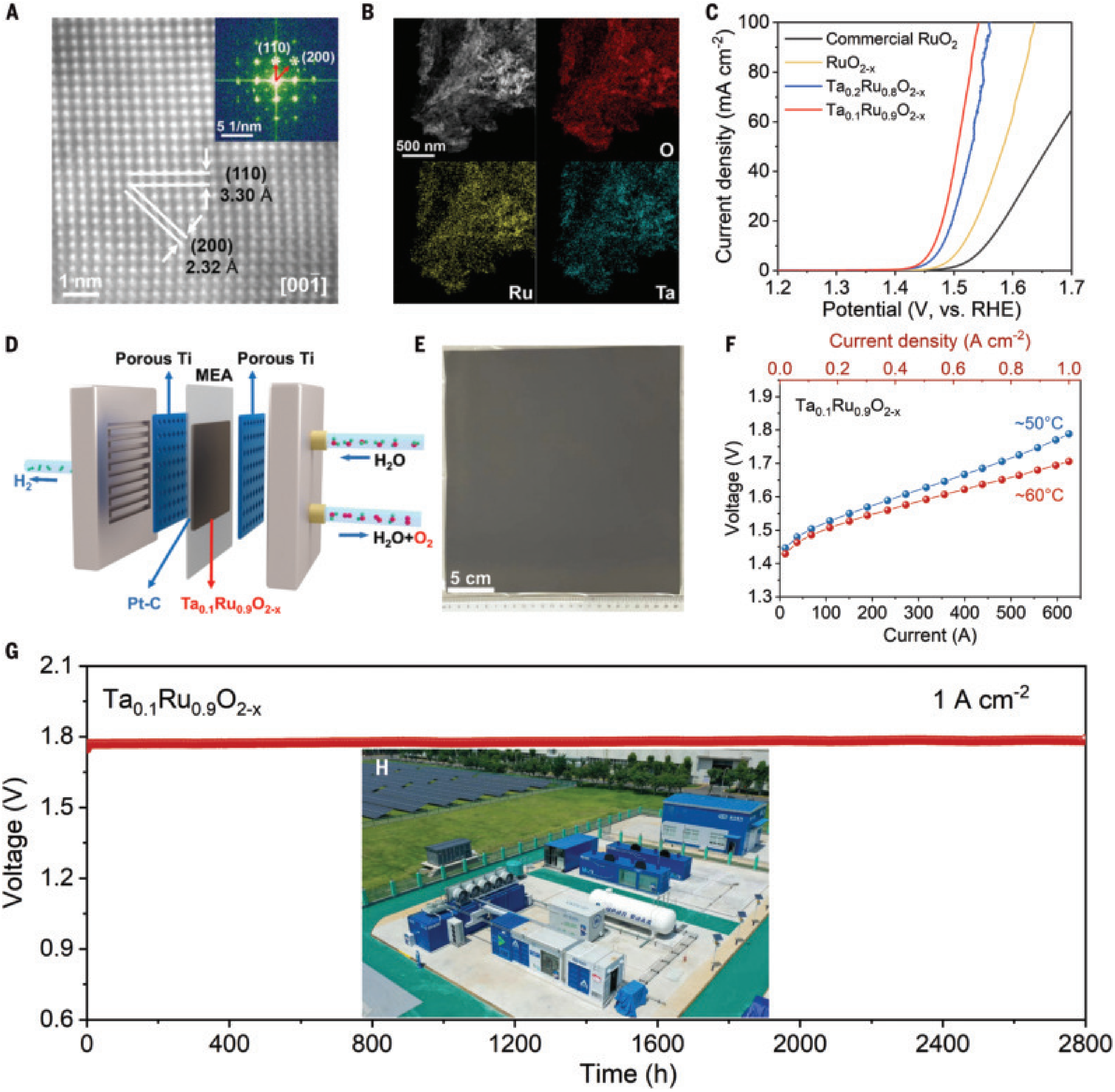


**Fig. 4. Nanosized $Ta_{0.1}Ru_{0.9}O_{2-x}$ electrocatalyst in industrial testing.** (**A**) HRTEM image (inset: fast Fourier transform pattern) and (**B**) EDS mapping of $Ta_{0.1}Ru_{0.9}O_{2-x}$. (**C**) OER curves of $Ta_{0.1}Ru_{0.9}O_{2-x}$, $Ta_{0.2}Ru_{0.8}O_{2-x}$, $RuO_{2-x}$, and commercial $RuO_2$ powder. (**D**) Illustration of a single-cell PEM-WE. (**E**) Photograph and (**F**) *I-V* curves of a 625-cm$^2$ MEA using $Ta_{0.1}Ru_{0.9}O_{2-x}$. (**G**) Voltage-time curves of PEM-WE with $Ta_{0.1}Ru_{0.9}O_{2-x}$. (**H**) A demonstration project in which solar-powered hydrogen production (of up to megawatt outputs) directly feeds a hydrogen refueling station.

tantalum chloride in ethanol solution was precipitated by alginate. The formed gel was subsequently dried and then calcined in air to produce ruthenium-based oxides (see materials and methods for details). Such synthesis could be readily scaled up to kilogram production, as we are currently doing to supply the commercialized Ta-$RuO_2$ catalyst products. The Ta content in Ta-$RuO_2$ is tunable, with the nanomaterials obtained denoted as $Ta_{\delta}Ru_{1-\delta}O_{2-x}$. As a control for comparison purposes, a $RuO_{2-x}$ sample was also synthesized by the same method but with no Ta added.

As confirmed by XRD (fig. S33A), doping $RuO_2$ with Ta did not change the rutile structure of $RuO_2$, at least for up to 20 at % Ta. As shown in the TEM image (fig. S33B), $Ta_{0.1}Ru_{0.9}O_{2-x}$ exhibited a porous and rough surface morphology, which provides a high surface area to make more OER active sites available. Figure 4A shows a high-resolution HRTEM image for a projection of the $Ta_{0.1}Ru_{0.9}O_{2-x}(001)$ plane in which the interplanar distances were 3.30 Å for (110) and 2.32 Å for (200), which is consistent with the XRD data. Element mapping enabled by EDS indicates that Ru, Ta, and O were homogeneously distributed in $Ta_{0.1}Ru_{0.9}O_{2-x}$ (Fig. 4B).

More characterizations can be found in figs. S34 to S38.

Figure 4C shows the results of OER evaluation on a rotating-disc electrode in 0.5 M $H_2SO_4$ electrolyte. From the OER polarization curves, $Ta_{0.1}Ru_{0.9}O_{2-x}$ shows a low overpotential of 226 mV at 10 mA $cm^{-2}$, appearing to have better OER activity performance over $Ta_{0.2}Ru_{0.8}O_{2-x}$ (fig. S39), synthesized $RuO_{2-x}$ (fig. S40A), and commercially available $RuO_2$ powder (fig. S40B). The Tafel plots for $Ta_{0.1}Ru_{0.9}O_{2-x}$ together with the benchmark samples are depicted in fig. S41A . The representative $Ta_{0.1}Ru_{0.9}O_{2-x}$ again outperformed other samples, exhibiting a low Tafel slope of 47.1 mV $dec^{-1}$, which is much smaller than that of commercial $RuO_2$. We evaluated the intrinsic activities of $Ta_{0.1}Ru_{0.9}O_{2-x}$ on the basis of TOFs at different overpotentials (fig. S41B). We found that the TOF of $Ta_{0.1}Ru_{0.9}O_{2-x}$ reached 0.086 $s^{-1}$ at an overpotential of 280 mV, which is the highest among the benchmark samples and 20 times higher than that of commercial $RuO_2$ (0.004 $s^{-1}$). The Nyquist plots shown as fig. S41C demonstrate that $Ta_{0.1}Ru_{0.9}O_{2-x}$ presented a smaller $R_{ct}$ compared with other samples, which implies that $Ta_{0.1}Ru_{0.9}O_{2-x}$ exhibited faster charge transfer toward OER. These results placed $Ta_{0.1}Ru_{0.9}O_{2-x}$ among the best catalysts in our controlled experiments.

To further illustrate the intrinsic OER activity for each electrocatalyst material, the apparent current density was also normalized to specific activity based on double-layer capacitance, $C_{dl}$ (fig. S42 and table S9). We found that $Ta_{0.1}Ru_{0.9}O_{2-x}$ and $Ta_{0.2}Ru_{0.8}O_{2-x}$ had similar specific activities, both outperforming $RuO_{2-x}$ and commercial $RuO_2$ (fig. S43). Notably, the $Ta_{0.1}Ru_{0.9}O_{2-x}$ powder catalyst exhibited a prominent $S$ number, which is within the same order of magnitude as that of $IrO_x$, showcasing the excellent stability of $Ta_{0.1}Ru_{0.9}O_{2-x}$ (fig. S44). We note that simply mixing $Ta_2O_5$ and $RuO_2$ led to only slightly improved OER activity (*29*) (fig. S45). We made similar observations for the thin film case (fig. S46) in which $Ta_2O_5$ was deposited on top of $RuO_2$ thin films through physical vapor deposition: This $Ta_2O_5$-modified $RuO_2$ (110) thin film had activity similar to that of $RuO_2$(110), and the drastic evolution during OER cycles was not suppressed.

We used $Ta_{0.1}Ru_{0.9}O_{2-x}$ electrocatalyst in MEAs (Fig. 4D). Current-voltage relation (*I-V*) curves demonstrate that MEAs with $Ta_{0.1}Ru_{0.9}O_{2-x}$ anode catalyst required 109 and 330 mV less in voltage to reach 1 A $cm^{-2}$ than did commercial $RuO_2$ and $IrO_2$ catalysts, respectively (fig. S47A), which implied that the MEA with $Ta_{0.1}Ru_{0.9}O_{2-x}$ had higher efficiency for water electrolysis. $Ta_{0.1}Ru_{0.9}O_{2-x}$ was further used as the anode catalyst to fabricate the industrial MEA as large as 25 cm by 25 cm (i.e., 625 $cm^2$ as shown in Fig. 4E). A single-cell water electrolyzer (fig. S48) enabled electrolyzing current as large as 625 A at voltage as low as 1.704 V (Fig. 4F), delivering a rated power of 1 kW and a hydrogen production rate at 0.523 normal cubic meters ($Nm^3$) $hour^{-1}$ (i.e., 145 ml $s^{-1}$ or 23.35 mol $hour^{-1}$).

During operation at a current density of 1 A $cm^{-2}$, $Ta_{0.1}Ru_{0.9}O_{2-x}$ exhibited a constant working voltage for at least 2800 hours at a decay rate <~14 μV $hour^{-1}$ (Fig. 4G), whereas $RuO_2$ lasted less than 50 hours before hitting the 2-V cutoff voltage (fig. S47B). This Ru-based electrocatalyst can be stably operated at a current density as high as 1 A $cm^{-2}$ in an MEA at industrial scale. As a demonstration project, the infrastructure for hydrogen production has been created and is integrated with solar-powered PEM-WE with outputs of up to megawatt rates, together with hydrogen storage and a hydrogen refueling station (Fig. 4H). Although the Ta-$RuO_2$ electrocatalyst is still being further evaluated in the industrial environment, we have no doubt that it will play an important part in the water electrolysis industry.

## Conclusions

We observed a structure-dependent corrosion on $RuO_2$ electrocatalysts during the OER process. Such corrosion is directly responsible for the poor stability (i.e., dissolution of Ru) of $RuO_2$ electrocatalysts. Ta doping suppresses the dissolution of Ru, while increasing the intrinsic activity of $RuO_2$ electrocatalysts toward OER. The industrial demonstration clearly shows that the poor stability issue of $RuO_2$ electrocatalysts has been successfully addressed and that Ru-based catalysts have more desirable OER activity than Ir-based catalysts, thus offering a highly attractive option for the electrocatalyst in PEM-WE.

## REFERENCES AND NOTES

1. V. R. Stamenkovic, D. Strmcnik, P. P. Lopes, N. M. Markovic, *Nat. Mater.* **16**, 57–69 (2016).
2. N. S. Lewis, D. G. Nocera, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15729–15735 (2006).
3. J. Song *et al.*, *Chem. Soc. Rev.* **49**, 2196–2214 (2020).
4. A. Buttler, H. Spliethoff, *Renew. Sustain. Energy Rev.* **82**, 2440–2454 (2018).
5. M. Carmo, D. L. Fritz, J. Mergel, D. Stolten, *Int. J. Hydrogen Energy* **38**, 4901–4934 (2013).
6. L. An *et al.*, *Adv. Mater.* **33**, e2006328 (2021).
7. K. E. Ayers *et al.*, *ECS Trans.* **33**, 3–15 (2010).
8. Y.-R. Zheng *et al.*, *Nat. Energy* **7**, 55–64 (2022).
9. Y. Chen *et al.*, *Sci. Adv.* **7**, eabk1788 (2021).
10. F. Cheng *et al.*, *Nat. Chem.* **3**, 79–84 (2011).
11. Y. Lin *et al.*, *Nat. Commun.* **10**, 162 (2019).
12. C. Liu *et al.*, *Nat. Catal.* **4**, 36–45 (2021).
13. Y. Qin *et al.*, *Nat. Commun.* **13**, 3784 (2022).
14. L. C. Seitz *et al.*, *Science* **353**, 1011–1014 (2016).
15. S. Hao *et al.*, *Nat. Commun.* **11**, 5368 (2020).
16. Z.-F. Huang *et al.*, *Nat. Energy* **4**, 329–338 (2019).
17. Y. Wen *et al.*, *J. Am. Chem. Soc.* **143**, 6482–6490 (2021).
18. Y. Yao *et al.*, *Nat. Catal.* **2**, 304–313 (2019).
19. B. Zhang *et al.*, *Science* **352**, 333–337 (2016).
20. L. Zhang *et al.*, *Adv. Mater.* **32**, e2002297 (2020).
21. P. C. K. Vesborg, T. F. Jaramillo, *RSC Adv.* **2**, 7933–7947 (2012).
22. J. Kibsgaard, I. Chorkendorff, *Nat. Energy* **4**, 430–433 (2019).
23. H. Over, *Chem. Rev.* **112**, 3356–3426 (2012).
24. M. Retuerto *et al.*, *Nat. Commun.* **10**, 2041 (2019).
25. S.-C. Sun *et al.*, *Angew. Chem. Int. Ed.* **61**, e202202519 (2022).
26. J. Wang *et al.*, *Chem* **8**, 1673–1687 (2022).
27. A. Zagalskaya, V. Alexandrov, *ACS Catal.* **10**, 3650–3657 (2020).
28. K. A. Stoerzinger *et al.*, *ACS Energy Lett.* **2**, 876–881 (2017).
29. V. Natarajan, S. Basu, *Int. J. Hydrogen Energy* **40**, 16702–16713 (2015).
30. Z. Yan *et al.*, *Electrochim. Acta* **157**, 345–350 (2015).
31. Y. Huang, R. J. Nielsen, W. A. Goddard 3rd, *J. Am. Chem. Soc.* **140**, 16773–16782 (2018).
32. L. D. Burke, O. J. Murphy, *J. Electroanal. Chem. Interfacial Electrochem.* **96**, 19–27 (1979).
33. K. A. Stoerzinger, L. Qiao, M. D. Biegalski, Y. Shao-Horn, *J. Phys. Chem. Lett.* **5**, 1636–1641 (2014).
34. K. Klyukin, A. Zagalskaya, V. Alexandrov, *J. Phys. Chem. C Nanomater. Interfaces* **123**, 22151–22157 (2019).
35. R. Kötz, S. Stucki, D. Scherson, D. M. Kolb, *J. Electroanal. Chem. Interfacial Electrochem.* **172**, 211–219 (1984).
36. H. C. Jo, K. M. Kim, H. Cheong, S.-H. Lee, S. K. Deb, *Electrochem. Solid-State Lett.* **8**, E39 (2005).
37. A. V. Korotcov, Y.-S. Huang, K.-K. Tiong, D.-S. Tsai, *J. Raman Spectrosc.* **38**, 737–749 (2007).
38. M. Naji *et al.*, *J. Raman Spectrosc.* **46**, 661–668 (2015).
39. C. Roy *et al.*, *ACS Energy Lett.* **3**, 2045–2051 (2018).
40. S. Geiger *et al.*, *Nat. Catal.* **1**, 508–515 (2018).
41. T. Weber *et al.*, *ACS Catal.* **9**, 6530–6539 (2019).
42. J. Kim, X. Yin, K.-C. Tsao, S. Fang, H. Yang, *J. Am. Chem. Soc.* **136**, 14646–14649 (2014).
43. W. A. Goddard III, S. Kwon, DFT Geometry and Energy, Version 1.0.0, Zenodo (2024); https://www.doi.org/10.5281/zenodo.14048116.

## ACKNOWLEDGMENTS

**Funding:** This work is primarily supported by Y.K.'s personal funds and resources, through Tianrui Technology Co., LTD. Computational work was supported by the Liquid Sunlight Alliance—which is supported by the US Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Fuels from Sunlight Hub under award no. DE-SC0021266 (W.A.G.) and by an individual fellowship from the Resnick Sustainability Institute at Caltech (S.K.)—and used Stampede3 at Texas Advanced Computing Center through allocation DMR160114 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by NSF grants 2138259, 2138286, 2138307, 2137603, and 2138296 (W.A.G.). This research used beamline 33-ID-D of the Advanced Photon Source, a DOE Office of Science user facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. This research used beamline 4-ID of the National Synchrotron Light Source II, a DOE Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under contract no. DE-SC0012704. **Author contributions:** Y.K. provided all the resources through Tianrui Tech and supervised the project. J.Z and Y.K. designed the experiment. J.Z. and X.F. performed the synthesis, characterization, and electrochemical tests. S.K. and W.A.G. carried out the theoretical calculations. K.C., L.Q., X.F., and Y.L. prepared the thin-films. X.L, F.X., J.W. and D.S. performed electron microscopy. J.Y. and H.S. facilitated the industrial tests. Y.W., T.U., Y.U., S.L., Y.L. and H.Z. are beam scientists who assisted in spectroscopy. G.C. performed the low energy electron microscopy. Y.K., J.Z., S.K., and W.A.G. drafted the manuscript. All authors discussed the results and wrote the paper. **Competing interests:** Y.K. through Tianrui Tech Co., LTD, commercialized the electrocatalysts described herein and holds the related patents. DongFang Boiler Group Co., LTD, is a megawatt water electrolyzer manufacturer and an energy solution provider. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials. The DFT data are available at Zenodo (*43*).

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ado9938
Materials and Methods
Supplementary Text
Figs. S1 to S48
Tables S1 to S13
Movie S1
References (*44*–*73*)

Submitted 5 March 2024; resubmitted 20 August 2024
Accepted 22 November 2024
10.1126/science.ado9938

CALORICS

# Organic ionic plastic crystals having colossal barocaloric effects for sustainable refrigeration

**Samantha L. Piper[1], Leena Melag[1]†, Mega Kar[1], Azra Sourjah[1], Xiong Xiao[2], Eric F. May[2], Kondo-Francois Aguey-Zinsou[3], Douglas R. MacFarlane[4]*, Jennifer M. Pringle[1]***

**Barocaloric (BC) materials offer the potential for highly energy-efficient refrigeration by generating heat absorption through the effect of pressure on a solid-solid phase transition. However, very few of the known materials have the required phase transition in the temperature regions necessary for domestic refrigeration or air conditioning. We introduce organic ionic plastic crystals (OIPCs) as a new family of BC materials. OIPCs display subambient transition temperatures, so-called “colossal” entropy changes (92 to 240 joules per kilogram per kelvin), and a high sensitivity to pressure, up to 23.7 kelvin per kilobar. The BC responses achieved with these prototype OIPC-BCs are tunable through structural modification of the ions; this wide matrix of possible combinations of structure and function indicates the scope of OIPCs as a new class of material for efficient and sustainable cooling technologies.**

Air conditioning and electric fans already account for 10% of all global electricity consumption, with projections indicating a massive increase over the next two decades in a warming climate, leading to a tripling of energy demand by 2050 to a level close to the total electricity consumption today of India and China combined (*1*). Adding to this burden, air conditioning will become a major contributor to peak electricity demand, exacerbating the issue of the intermittency of renewables and the need for large-scale energy storage, all the while emitting more heat-trapping gases (the global-warming potential of common commercial refrigerant R-134a is ~1300 times greater than that of $CO_2$) (*2*, *3*). Advances in cooling technologies are therefore seen as an urgent priority in reducing this burgeoning energy cost, with the International Energy Agency’s Net Zero Emissions by 2050 Scenario calling for a 50% increase in air conditioning efficiency by 2030 (*4*).

Solid-state caloric materials are emerging as alternatives to the refrigerants currently used in conventional cooling technologies (*5*), which despite steady improvements in efficiency have remained largely unchanged for more than a century. These materials exhibit a thermal response to a perturbation such as a pressure change; their associated cooling systems can therefore operate on the same thermodynamic principles as the vapor-compression cycle, while evading the use of high global-warming-potential working fluids and offering the potential of more energy-efficient refrigeration. Caloric materials reported in the literature use various external perturbations to drive the process, including magnetic (*6*), electric (*7*), and mechanical (*8*). The current drawback of these systems is the typically large fields required to induce a reversible thermal response. One of the most promising options, the barocaloric effect, uses hydrostatic pressure to drive a phase change in the solid material; the first reports of this appeared toward the end of the 1990s (*9*). The subsequent interest eventually led to the discovery of so called “colossal” entropy changes ($\Delta S_{\text{s-s}} > 100$ J kg$^{-1}$ K$^{-1}$) occurring at the solid-solid (s-s) transition in neopentane-derived plastic crystals (*10*, *11*), approaching the performance of modern fluid refrigerants and vastly outperforming other field-driven caloric effects. Although accessible at moderate pressures because of the high sensitivity of the transition temperature ($T_{\text{s-s}}$) to pressure (a large value of $dT_{\text{s-s}}/dp$), the practical applicability of these materials in cooling applications is hindered by their above-ambient transition temperatures (for example, neopentyl glycol, $T_{\text{s-s}}$ = 41°C); typically the transition temperature needs to sit close to but below the cooling temperature required. Colossal barocaloric effects have since been observed in other families of materials—including spin-crossover complexes (*12*), inorganic-organic hybrid perovskites (*13*), and alkylammonium halides (*14*, *15*)—although various drawbacks such as high driving pressures and, critically, inappropriately high (typically above ambient) transition temperatures are still to be overcome.

One of the original reasons for interest in solid-state barocaloric cycles is the urgent need to replace hydrofluorocarbons as refrigerants because of their high global warming potential. The technology also has the potential to improve the energy cost of various cooling and refrigeration cycles—if high-performance barocaloric materials having $T_{\text{s-s}}$ in the appropriate range for each application can be discovered (*16*). Detailed analyses of the refrigeration cycles that can be applied to such materials show that under certain conditions, the barocaloric cycle can substantially outperform traditional vapor compression cycles in terms of energy cost in domestic refrigeration and air conditioning (*17*, *18*). We describe a number of such materials that are promising in their own right and that also illustrate the potential of a large class of compounds, organic ionic plastic crystals (OIPCs) (Fig. 1A), that have not previously been recognized in this context. OIPCs possess the same advantageous features as those of molecular plastic crystals, with the additional advantage of having negligible vapor pressure (*19*). Unlike molecular plastic crystals, OIPCs consist of cations and anions whose dynamics change in complex ways at the solid-solid transition (*20*, *21*). Their thermal properties are therefore diverse; multiple solid phases are sometimes observed (*20*), whereas other OIPCs show only one high-entropy solid-solid phase transition (*22*). Those possessing a single, major first-order transition—which is associated with a substantial increase in dynamic freedom, symmetry, and therefore entropy (Fig. 1B)—are preferred for barocaloric applications (*23*). The reversible, first-order solid-solid transitions in OIPCs often occur at the subambient temperatures suitable for cooling applications, making them highly attractive for energy-efficient cooling technologies.

A broad range of organic cations and anions have been reported to form OIPCs (fig. S1) (*24*–*28*); selection of appropriate ions is therefore guided by the suitability of the ion properties (such as chemical stability) for the intended application. For example, the inclusion of the bis(fluorosulfonyl)imide ([FSI]$^-$) or bis(trifluoromethanesulfonyl)imide ([TFSI]$^-$) anions (chemical structures are provided in Fig. 1A) in OIPCs has been shown to produce quite distinct reorientational dynamics—often involving rotation between the transoid and cisoid conformations in the plastic phases—that contribute to large entropy changes at the solid-solid transitions (*29*–*31*). Lattice vibrations can also contribute to the phase transition entropy change, quite substantially in some cases (*32*). When these anions are coupled with *N*-methyl-*N*-isopropylpyrrolidinium ([C$_{i3}$mpyr]$^+$) and similar cations (Fig. 1A), the resulting OIPCs display solid-solid transitions at low temperatures. We report the first barocaloric study of four prototype OIPCs (Fig. 1A), which we chose on the basis of their known thermophysical properties and transition temperatures, including the observation of large entropy changes in our previous work (table S6) (*22*, *33*). The nature of the phase changes within these OIPCs also suggests the possibility of large volume changes at the transition, leading to the prediction of high sensitivity of the transition temperatures

---

[1]Institute for Frontier Materials, Deakin University, Burwood, VIC, Australia. [2]Fluid Science and Resources Division, Department of Chemical Engineering, The University of Western Australia, Crawley, WA, Australia. [3]MERLin, School of Chemistry, University of Sydney, Sydney, NSW, Australia. [4]School of Chemistry, Monash University, Clayton, VIC, Australia.

***Corresponding author. Email: jenny.pringle@deakin.edu.au (J.M.P.); douglas.macfarlane@monash.edu (D.R.M.)**

†Present address: Process Development, CSIRO Mineral Resources, Clayton, VIC, Australia.

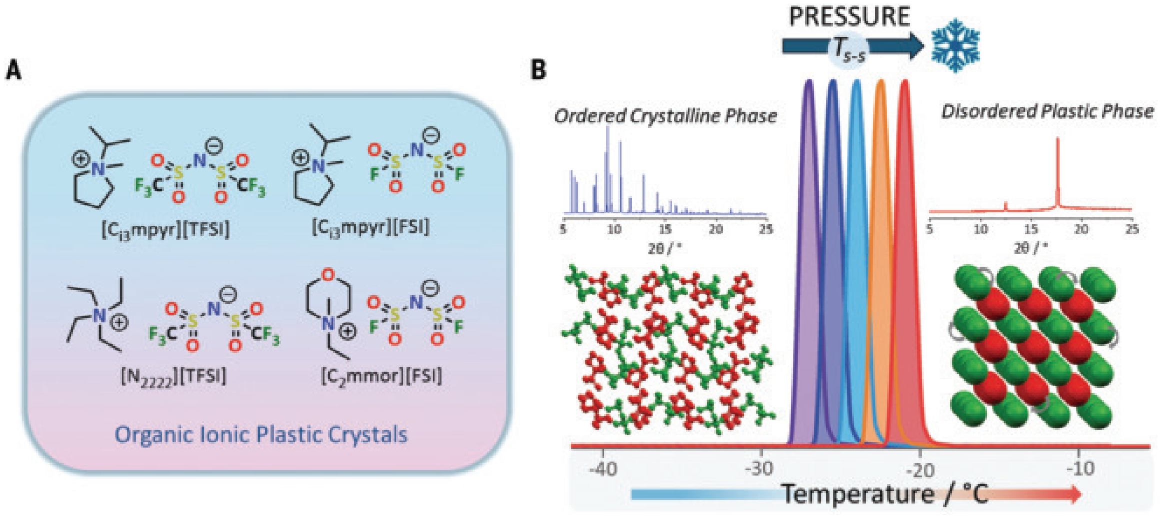


**Fig. 1. Organic ionic plastic crystals as barocaloric materials.** (**A**) Chemical structures of the OIPCs studied. [C$_{i3}$mpyr], *N*-methyl-*N*-isopropylpyrrolidinium; [N$_{2222}$], tetraethylammonium; [C$_2$mmor], 4-ethyl-4-methylmorpholinium; [TFSI], bis(trifluoromethanesulfonyl)imide; [FSI], bis(fluorosulfonyl)imide. (**B**) A typical DSC trace showing the dominant solid-solid endothermic phase transition in [C$_{i3}$mpyr][FSI] in which the static, low-symmetry, crystalline phase transitions into a phase in which isotropic tumbling of the ions on their lattice sites leads to a higher-symmetry crystal structure at the transition temperature $T_{s\text{-}s}$. The different colored endotherms demonstrate the shift of $T_{s\text{-}s}$ to higher temperatures at increased applied hydrostatic pressure; the magnitude of this shift is $dT_{s\text{-}s}/dp$. (Inset) Measured powder patterns (*54*) and low-temperature crystal structure for [C$_{i3}$mpyr][FSI]. The high symmetry structure is an artist's impression based on the XRD.

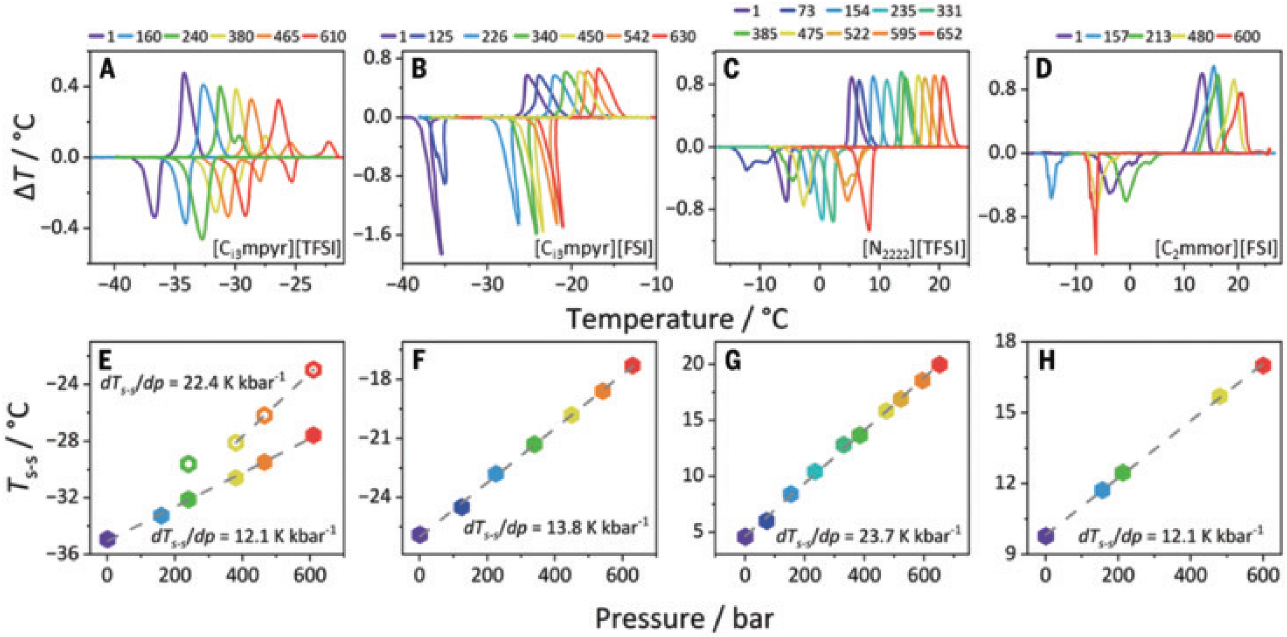


**Fig. 2. High-pressure differential thermal analysis.** (**A** to **D**) HP-DTA curves of (A) [C$_{i3}$mpyr][TFSI], (B) [C$_{i3}$mpyr][FSI], (C) [N$_{2222}$][TFSI], and (D) [C$_2$mmor][FSI] at various hydrostatic pressures (shown in the color legends where the unit is bar). (**E** to **H**) Onset temperature of the endothermic transitions as a function of pressure, in which the gradient of the linear fit defines the sensitivity of the transition temperatures to pressure ($dT_{s\text{-}s}/dp$). Because of the splitting of the peaks at higher pressures for [C$_{i3}$mpyr][TFSI], both components are plotted in (E) showing that the sensitivity of the smaller component (open symbols) to pressure is roughly double that of the main component (solid symbols) (22.4 versus 12.1 K kbar$^{-1}$). Error values derived from the linear fits are shown in Table 1.

to pressure. We used differential scanning calorimetry (DSC) and high-pressure differential thermal analysis (HP-DTA) to directly investigate the effect of pressure on the transitions, along with pycnometry to study the volumetric changes and allow an analysis of the barocaloric properties. Where possible, we also used single-crystal x-ray crystallography to probe the crystal structure changes across the solid-solid transition.

The OIPCs we present represent only a small number of examples of a potentially enormous family of materials that may offer pressure-sensitive, colossal barocaloric effects at the low temperatures suitable for cooling applications. The extraordinary scope for tuning these properties, through modifications to the myriad possible cation and anion structures, offers

**Table 1. Phase transition and barocaloric properties.** $T_{s\text{-}s}$ was determined by the onset of the transition peak observed upon warming by means of differential scanning calorimetry at 1 bar, by using a scan rate of 1°C $min^{-1}$ and a sample size of 2 to 5 mg. $dT_{s\text{-}s}/dp$ was determined by using warming data from high-pressure differential thermal analysis at different hydrostatic pressures between 1 and 700 bar (Fig. 2), by using a scan rate of 1°C $min^{-1}$ and sample size of 40 to 80 mg. Uncertainties in $dT_{s\text{-}s}/dp$ are derived from the linear fit of the $T_{s\text{-}s}$ versus $p$ data. Reported values of $\Delta v_{s\text{-}s}$ were determined by using the solvent displacement pycnometry method described in the supplementary materials, for all compounds except for $[C_{i3}mpyr][TFSI]$, which because of its very low $T_{s\text{-}s}$ was instead determined with XRD methods.

| Material | $T_{s\text{-}s}$ (°C) (±1°C) | $\Delta H_{s\text{-}s}$ (J $g^{-1}$) (±10%) | $dT_{s\text{-}s}/dp$ (K $kbar^{-1}$) | $\Delta S_{s\text{-}s}$ (J $kg^{-1}$ $K^{-1}$) (±10%) | $\Delta v_{s\text{-}s}$ (%) (±10%) |
|---|---|---|---|---|---|
| $[C_{i3}mpyr][TFSI]$* | –37 | 22 | 12.1 (±0.3) | 92 | 2.5 |
| $[C_{i3}mpyr][FSI]$ | –28 | 47 | 13.8 (±0.2) | 190 | 3.8 |
| $[N_{2222}][TFSI]$ | 4 | 45 | 23.7 (±0.3) | 160 | 6.2 |
| $[C_2mmor][FSI]$ | 10 | 68 | 12.1 (±0.2) | 240 | 4.6 |

*Because the transition splits into two peaks for $[C_{i3}mpyr][TFSI]$ at pressures >240 bar, the value of $dT_{s\text{-}s}/dp$ reported in the table represents that of the dominant peak. The $dT_{s\text{-}s}/dp$ of the second, smaller component, which becomes linear above 380 bar, is 22.4 (±0.1) K $kbar^{-1}$.

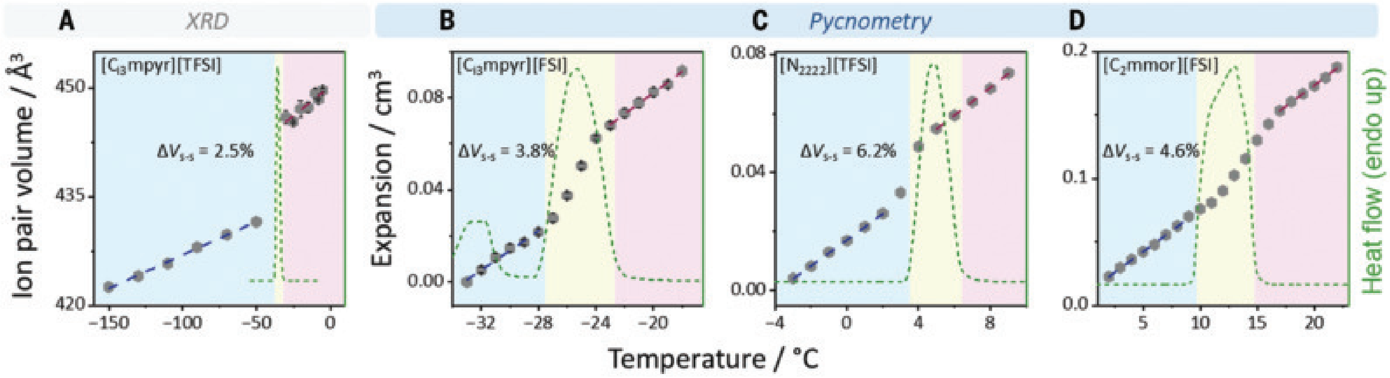


**Fig. 3. Volume change determination.** (**A**) One-half the unit cell volume of $[C_{i3}mpyr][TFSI]$ ($T < -40°C$) and one-fourth the unit cell volume ($T > -40°C$)—that is, the volume of an ion pair. $\Delta V_{s\text{-}s}$ is determined by the difference in the volume predicted by the linear fits at the temperature of the transition. (**B** to **D**) Solvent displacement data for 0.8867 g $[C_{i3}mpyr][FSI]$, 0.3754 g $[N_{2222}][TFSI]$, and 1.0228 g $[C_2mmor][FSI]$, in which the linear fits of solvent expulsion before and after the transition temperatures were used to determine the volume of solvent ejected because of the transition: $\Delta V_{s\text{-}s}$. All data points indicate the average of at least duplicates, and error bars are included when the number of replicates is three. Volume data are superimposed on the DSC trace of the materials to show the correlation with phase transitions.

great potential to maximize the barocaloric performance of this class of materials.

## Results and discussion

The HP-DTA curves of the four OIPCs at various pressures are shown in Fig. 2. The transition temperatures of the materials ($T_{s\text{-}s}$) are all subambient and range from –37°C to 10°C (Table 1). The influence of hydrostatic pressure on the thermal properties allows some of the barocaloric properties of the materials to be directly derived, including the sensitivity of the transition temperature to pressure ($dT_{s\text{-}s}/dp$) (Table 1), and its positive gradient indicates a conventional barocaloric effect. These values are comparable with example compounds from the literature that have been studied for barocaloric applications and are considered to be some of the top performing materials (table S19).

The $dT_{s\text{-}s}/dp$ is above 10 K $kbar^{-1}$ for all of the materials and is as high as 23.7 (±0.3) K $kbar^{-1}$ for $[N_{2222}][TFSI]$, which is among the highest reported values of $dT_{s\text{-}s}/dp$ for barocaloric materials (literature comparisons are available in table S19). This value is a key parameter because it dictates the pressure change required to achieve the phase transition at any given temperature. However, finding BC materials with simultaneously large values of $dT_{s\text{-}s}/dp$ and $\Delta S_{s\text{-}s}$ can be challenging because of the inverse relationship between these two properties indicated by the Clausius-Clapeyron (C-C) relationship. A rearranged form of the C-C equation, $\left(\frac{dT_{s-s}}{dp}\right) = \frac{\Delta v_{s-s}}{\Delta S_{s-s}}$, suggests that achieving a large value of both $dT_{s\text{-}s}/dp$ and $\Delta S_{s\text{-}s}$ requires a large volume change ($\Delta v_{s\text{-}s}$) at the transition. We used a pycnometer-based method for measuring $\Delta V_{s\text{-}s}$ for OIPCs with $T_{s\text{-}s} > -30°C$ and used variable-temperature x-ray diffraction (XRD) for the OIPC with the lowest $T_{s\text{-}s}$. Our results indicate volume changes on the order of $1 \times 10^{-5}$ to $5 \times 10^{-5}$ $m^3$ $kg^{-1}$, equating to volume changes between 2 and 7% (Fig. 3, Table 1, and table S7). This facile method of measuring $\Delta V_{s\text{-}s}$ is also useful because it can provide an early estimate of $dT_{s\text{-}s}/dp$ from 1 bar data and the C-C equation, eliminating the need for complex or expensive high-pressure equipment to screen barocaloric properties of new materials [(34), section 1.3].

The transition peaks of $[C_{i3}mpyr][TFSI]$ appear to split at higher pressures, which is behavior that has also been observed for other materials (13, 35). In this case, the higher $dT_{s\text{-}s}/dp$ of the second, smaller component may indicate that there are two transitions occurring in very close succession at 1 bar that become differentiable at higher pressures because of their varying sensitivity to pressure. Nevertheless, the OIPC demonstrates very stable cycling behavior, with no detectable changes to the temperature or size of the transition after 100 heating-cooling cycles at atmospheric pressure (fig. S13).

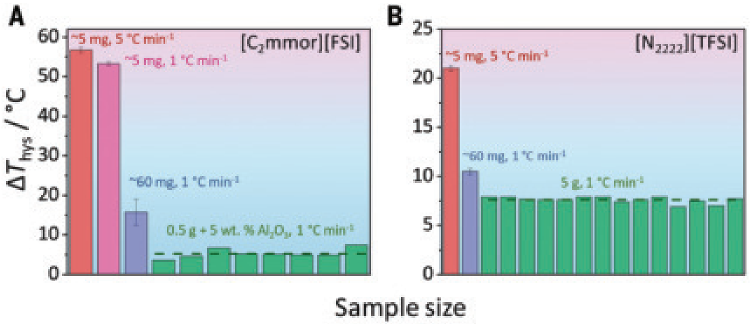


**Fig. 4. Influence of experimental parameters on $\Delta T_{hys}$.** (**A** and **B**) Values of $\Delta T_{hys}$ observed under different conditions for (A) [$C_2$mmor][FSI] and (B) [$N_{2222}$][TFSI]. The red bars indicate the $\Delta T_{hys}$ determined from DSC experiments (ramp rate, 5°C min$^{-1}$) where the sample mass is ~5 mg. In (A), the pink bar indicates $\Delta T_{hys}$ determined from DSC experiments with a sample size of 5 mg and reduced cooling rate of 1°C min$^{-1}$. The purple bars indicate the average $\Delta T_{hys}$ from high-pressure DTA experiments (ramp rate, 1°C min$^{-1}$) in which the sample mass was ~60 mg. The green bars indicate values of $\Delta T_{hys}$ for consecutive DTA cooling cycles of a sample (A) with or (B) without an added nucleating agent $Al_2O_3$ (5 wt %) (ramp rate, 1°C min$^{-1}$). Average values of $\Delta T_{hys}$ are indicated with the dashed green lines.

### Solid-state structure and dynamics

The structure and dynamics of OIPCs in their plastic and low-temperature phases have been probed by means of techniques such as solid-state nuclear magnetic resonance spectroscopy (*31*, *36*), XRD (*31*), and MD simulations (*21*). These have revealed the evolution of different molecular motions over the solid-solid transition, often including rotational or reorientational motions as well as increased diffusional motions of the cations and/or anions. The free volume in the plastic phase is an intrinsic material property, evidence for which has been provided with positron annihilation lifetime spectroscopy (*37*). Free volume has been well studied as the basis for ionic conduction (*38*), in the usual application of OIPCs as electrolytes, and we show that free volume is also highly beneficial for barocaloric performance.

X-ray powder or single-crystal diffraction are the ideal probes of structure (*20*, *39*, *40*), but resolving atomic positions in the plastic phase is typically very challenging because of the substantial motional dynamics present. In the powder patterns, the plastic phases of OIPCs characteristically display very few diffraction peaks because of the short-range molecular motions (rotations) that produce disorder in some of the reflections (*29*). The powder patterns in phase I (the most plastic phase) of the OIPCs show that those with the highest values of $\Delta S_{s\text{-}s}$ show the simplest diffraction patterns, with only two diffraction peaks in the patterns of [4-ethyl-4-methylmorpholinium ([$C_2$mmor])[FSI] and [$C_{i3}$mpyr][FSI] at low values of 2θ (fig. S12, B and D). This is supported by the polarized optical microscopy images (fig. S14) that only show birefringence for the TFSI OIPCs, indicating a simpler cubic structure for the FSI salts.

Although resolving single-crystal structures of plastic phases is inherently challenging, a structure of the supercooled plastic phase of [$N_{2222}$][TFSI] has been reported previously (fig. S11) (*41*). Half of all cations and anions are disordered in the structure, which indicates that the dynamics of both ions contribute to the high $\Delta S_{s\text{-}s}$ of this salt. The [$N_{2222}$]$^+$ cation appears to dictate the extent of disorder in the structure, with the [TFSI]$^-$ anions "flexing" ($C_2 \leftrightarrow C_2$ disordering mode) in the space created by neighboring cations (*41*).

Demonstrating the influence that both ions can have on the phase transition properties of OIPCs, the lower-temperature structures of [$C_{i3}$mpyr][TFSI] and [$C_{i3}$mpyr][FSI] both solve as relatively static ion pair assemblies with reasonable anisotropic displacement parameters (figs. S7 to S10). Despite their similar three-dimensional packing arrangements and mutual cation, the $\Delta S_{s\text{-}s}$ of [$C_{i3}$mpyr][FSI] is approximately double that of [$C_{i3}$mpyr][TFSI]. The greater extent of disorder in the FSI analog, possibly owing to its smaller size and ease of multi-axis rotation, results in fewer diffraction peaks in its phase I powder pattern compared with that of [$C_{i3}$mpyr][TFSI] (fig. S12, A and B). Clearly, both ions have a crucial role to play in determining the nature and extent of disorder in the plastic phase and thus the key properties of the phase transition. This highlights the broad scope of prospective OIPCs that could be studied for this application and the vast potential to optimize their barocaloric properties through structural optimization of the constituent ions.

### Reversible barocaloric effects: $p_{rev}$

The key parameters that dictate the applicability of any barocaloric material are considered to be (i) a transition temperature appropriate to the application (subambient for most cooling applications) and the sensitivity of the transition temperature to pressure ($dT_{s\text{-}s}/dp$); (ii) the pressure change required to cause the phase change to reversibly occur, $p_{rev}$ (this is a particularly relevant metric in regard to the energy cost of the refrigeration cycle because it dictates the minimum work required to drive the cycle); and (iii) the size of entropy change that can be reversibly achieved in a pressure cycle (*42*). We further analyzed the latter two.

OIPCs that exhibit a large entropy change during the solid-solid, disordering phase transition upon warming may require initiation of "reordering" during the reverse transition on cooling, potentially requiring large supercooling, similar to that observed in many materials in the transition from liquid to solid (*43*). This temperature hysteresis ($\Delta T_{hys}$) is experimentally defined by the temperature difference between the onset of the transition on warming and the onset of the reverse transition on cooling. This is a pivotal property of a barocaloric material. Together with $[dT_{s\text{-}s}/dp]_{cooling}$, $\Delta T_{hys}$ provides an estimate of the minimum applied pressure ($p_{rev}$) that is required to bring about the reversible change of phase: $p_{rev} = \Delta T_{hys}/[dT_{s\text{-}s}/dp]_{cooling}$ (*44*). This value needs to be as small as possible to minimize both equipment costs as well as the energy costs associated with cyclic pressure generation.

Assuming that the starting pressure is 1 bar, the $p_{rev}$ values we estimated for the OIPCs are shown in table S19. The low $\Delta T_{hys}$ in [$C_{i3}$mpyr][TFSI], combined with a relatively large $dT_{s\text{-}s}/dp$, results in an extremely low $p_{rev}$ of just 42 bar for this material; this value is among the lowest according to our comparison with recent literature reports across a variety of barocaloric materials (table S19).

The initiation of the ordered phase during cooling may be related to interactions at the surface of the material (*45*, *46*) and is often observed to be a function of the volume of the sample. Therefore, practically relevant estimates of $p_{rev}$ must be based on volumes of sample larger than is typical for DSC or DTA measurements. Other factors that influence $\Delta T_{hys}$ include (i) the rate of cooling, with slower rates usually decreasing $\Delta T_{hys}$, and (ii) the presence of solid impurities or additives (*47*). An approach to reducing $\Delta T_{hys}$ thus involves the addition of nucleating agents to a sample to increase the number of nucleation sites per unit volume. This method is often successfully used in the liquid-solid phase change materials used for thermal energy storage (*48*, *49*).

To investigate these strategies to reduce $\Delta T_{hys}$, we studied the order-disorder phase transition in [$C_2$mmor][FSI], which initially exhibits the largest DSC $\Delta T_{hys}$ (~55°C) of our materials. Over eight consecutive heating and cooling cycles of a larger sample and/or with an added nucleating agent, aluminum oxide ($Al_2O_3$), the average crystallization temperature decreased to 4.7°C (SD = 1.1°C) (Fig. 4A and figs. S15 to S17), demonstrating a decrease in $\Delta T_{hys}$ of ~50°C compared with the data obtained

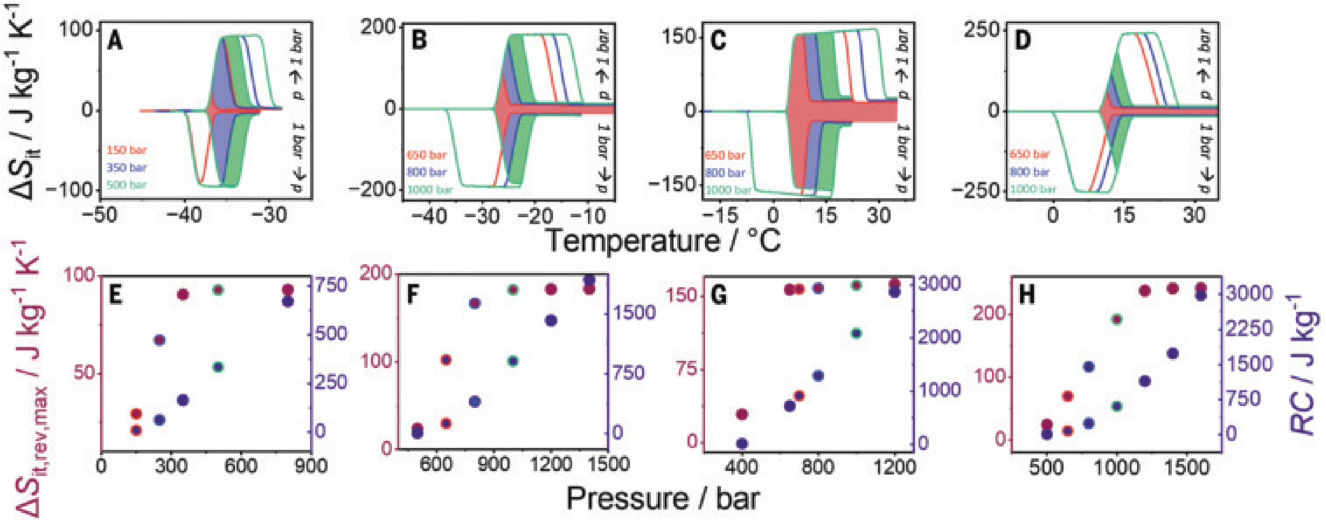


**Fig. 5. Barocaloric performance metrics.** (**A** to **D**) Isothermal entropy curves (solid lines) of (A) [C$_{i3}$mpyr][TFSI], (B) [C$_{i3}$mpyr][FSI], (C) [N$_{2222}$][TFSI], and (D) [C$_2$mmor][FSI] upon compression ($\Delta S_{it} < 0$ J kg$^{-1}$ K$^{-1}$) and decompression ($\Delta S_{it} > 0$ J kg$^{-1}$ K$^{-1}$) for select pressure changes denoted in the color key. The shaded areas indicate the overlap between the compression curve and the decompression curve when reflected over the *x* axis; these indicate the reversible entropy change ($\Delta S_{it,rev}$) and the temperature span over which a reversible entropy change can be accessed for a change from 1 bar to a given pressure (additional to the area shaded for lower pressure changes). (**E** to **H**) From the curves in (A) to (D), values of $\Delta S_{it,rev,max}$ for each pressure change (and some additional pressure changes) are plotted in (E) to (H), along with corresponding values of refrigeration capacity, *RC*. The symbols for data extracted from the curves illustrated in (A) to (D) are highlight-colored in (E) to (H) to match the corresponding pressure. *RC* values are indicative of the cooling power of a given refrigerant under a defined pressure cycle (*11*). As $\Delta T_{FWHM}$ increases continuously with increasing pressure, *RC* increases beyond the point at which $\Delta S_{it,rev,max}$ is reached (for example, 350 bar for [C$_{i3}$mpyr][TFSI]). Although $\Delta T_{hys}$ plays a substantial role in determining the value of *RC* at low pressures, this role is reduced at higher pressures, at which $\Delta S_{it,rev,max}$ and $dT_{s\text{-}s}/dp$ become the more relevant parameters. Hence, whereas (E) the *RC* of [C$_{i3}$mpyr][TFSI] is notably high at low pressures because of its low $\Delta T_{hys}$ (163 J kg$^{-1}$ at 350 bar), (G) [N$_{2222}$][TFSI] has the highest *RC* amongst these OIPCs at increased pressures, reaching 1292 J kg$^{-1}$ at a still moderate applied pressure of 800 bar.

from DSC experiments (table S15). Similarly, a 5-g sample of [N$_{2222}$][TFSI] crystallized at an average temperature of −4.2°C (SD = 0.25°C) over 15 cycles (Fig. 4B, table S16, and fig. S19). By comparison, the smaller sample size used in DSC experiments resulted in crystallization at −17°C. Hence, we could achieve a substantial reduction in $\Delta T_{hys}$ by using these relatively simple approaches to reducing supercooling.

With these effects in mind, we used the lowest observed values of $\Delta T_{hys}$ to estimate values of $p_{rev}$ (table S19), aware that further improvements in these properties will appear in the larger quantities of material used in real applications. Practically, there will always be thermal conduction limitations, full allowance for which involves the design parameters of the practical device and are beyond the scope of this paper. Typical thermal conductivity values for organic materials such as these are on the order of 0.1 to 0.3 W m$^{-1}$ K$^{-1}$ (*50*) and can be enhanced by the inclusion of materials such as carbon or graphene (*51*). A serendipitous side effect of the addition of such materials is likely to be a further decrease in $\Delta T_{hys}$ from nucleation effects.

### Evaluation of barocaloric effects

We can analyze the barocaloric properties in further detail using the procedures described by Li (*10*) and Lloveras (*11*). The calorimetry data were first transformed into entropy as a function of temperature at each pressure [(*34*), section 2.4], and we show an example of these isobaric $S_{ib}(T, p)$ plots in fig. S21A.

The isothermal entropy changes ($\Delta S_{it}$) relevant to a pressure cycle were then calculated at any temperature from the differences in these curves. Bearing in mind that elevating the pressure from $p_0 \rightarrow p_1$ causes the formation of the ordered (lower volume) phase

$$\Delta S_{it}\ (p_0 \rightarrow p_1) = \Delta S_{ib(cooling)}(T, p_1) - \Delta S_{ib(cooling)}(T, p_0)$$

And on decompression from $p_1 \rightarrow p_0$, which causes the formation of the disordered phase

$$\Delta S_{it}\ (p_1 \rightarrow p_0) = \Delta S_{ib(heating)}(T, p_0) - \Delta S_{ib(heating)}(T, p_1)$$

We show these curves for the OIPCs at several pressures assuming that $p_0$ is 1 bar (Fig. 5, A to D).

The maximum entropy change that can be reversibly achieved in a pressure cycle at any temperature ($\Delta S_{it,rev}$) is represented by the overlap of the compression and decompression curves. We calculated this as shown in the example in fig. S21B, with the overlap ($\Delta S_{it,rev}$) regions shown as solid shading. The peak of this overlap region indicates the temperature at which the entropy change is maximized ($\Delta S_{it,max,rev}$). Carrying out this calculation across a range of pressures shows that $\Delta S_{it,max,rev}$ increases with the size of the pressure change up to a limiting value (Fig. 5, E to H), which is approximately that observed in the DSC thermograms. For example, our plots for [C$_{i3}$mpyr][TFSI] indicate that the entirety of the entropy associated with the transition can be reversibly achieved at $p_1 \geq 350$ bar (Fig. 5, A and E). The refrigeration capacity (*RC*) can also be calculated (Fig. 5, E to H) from $RC = \Delta S_{it,rev,max} \Delta T_{FWHM}$, where $\Delta T_{FWHM}$ is the peak width at half maximum of the $\Delta S_{it,rev}$ curves (an example is provided in fig. S23).

The upward displacement of this pressure from $p_{rev}$ highlights the role of the endothermic and exothermic DSC peak widths in achieving the optimum entropy change at low pressures. The nature of the measurement means that we cannot observe the fundamentally sharp step in entropy during an ideal first-order, isothermal phase transition. The widths are at least partly due to thermal conduction limitations in these unoptimized experiments (poor contact with sample container and finely divided samples) (*13*), and therefore, to the extent that the transitions are first order, the barocaloric performance indicated in Fig. 5 is likely an underestimate of the fundamental material performance. An additional contribution to $\Delta S_{it}$ arising from compression of the material can be as large as 20% of the phase transition entropy change (fig. S23A) [(*34*), section 2.6].

The *RC* performance parameter can be used to compare barocaloric materials generally by calculating the *RC* that can be achieved for a given pressure change ($RC/\Delta p$). Values of $RC/\Delta p$ calculated for these OIPCs at various pressures can be compared with some leading examples from the literature (table S19). The high values of $\Delta S_{\text{s-s}}$ and $dT_{\text{s-s}}/dp$, in particular for $[N_{2222}]$[TFSI], result in a rapid increase in $RC/\Delta p$ at pressures > $p_{\text{rev}}$ with this material. At 1 kbar, this pressure-normalized value of *RC* of $[N_{2222}]$[TFSI], at 2.1 J $\text{kg}^{-1}$ $\text{bar}^{-1}$, is comparable with that of other high-performance materials.

Under adiabatic conditions, the maximum temperature change generated by the transition can be also estimated (*52*, *53*) from $\Delta T_{\text{ad,max}} = T\Delta S_{\text{ib}}/C_{\text{p}}$, for which we obtained values (table S18) as high as 62°C for $[C_2\text{mmor}]$[FSI] and 42°C for $[N_{2222}]$[TFSI]. These are among the highest reported in the literature, including those for $NH_4I$ at 34°C (*52*) and $(DA)_2MnCl_4$ at 45°C (*53*). This represents an upper bound on $\Delta T_{\text{ad}}$; a more practically useful estimate of the accessible adiabatic temperature change for a given pressure change can be obtained by the method described in (*34*), section 2.5. This "quasi-direct" calculation yields values of 23°C for $[N_{2222}]$[TFSI] and 11°C for $[C_2\text{mmor}]$[FSI] at 1 kbar (fig. S22 and table S18).

Hence, these OIPCs are not only some of the few examples of barocaloric materials with transition temperatures in the appropriate range for cooling applications, they are also among the highest-performing materials by some of these metrics. For the OIPCs that show substantial hysteresis ($[C_{i3}\text{mpyr}]$[FSI], $[C_{i3}\text{mpyr}]$[TFSI], and $[N_{2222}]$[TFSI]), *RC* could likely be improved at substantially lower pressures with the larger sample sizes relevant to application.

## Conclusions

OIPCs represent a highly promising family of barocaloric materials that can undergo high-entropy transitions upon the application of moderate hydrostatic pressures. They offer solid-solid transitions at the subambient temperatures required for most cooling applications. The very low transition temperatures in some OIPCs, including two reported here ($[C_{i3}\text{mpyr}]$[TFSI] and $[C_{i3}\text{mpyr}]$[FSI]), enable a distinct scope to further tailor the transition temperature to suit a given application by increasing the lower operating pressure, $p_0$, which has otherwise been taken as 1 bar in this work. However, this scope is ultimately limited by the increased work requirements of such a system.

A distinguishing feature of these OIPCs is the isothermal nature of their first-order, solid-solid phase transitions; in the absence of thermal conduction limitations, this would allow the entire entropy associated with the phase transitions to be accessed at $p_{\text{rev}}$. We have shown that values of $\Delta T_{\text{hys}}$ are highly sensitive to sample volumes and are reduced in larger sample sizes, as in real application, and can be further reduced by the inclusion of additives. Thus, there are simple strategies that could further improve the barocaloric performance metrics of OIPCs in an optimized device.

The ability to modify the chemical structure of both substituent ions of an OIPC to tailor phase-transition properties provides enormous scope for optimizing barocaloric properties. As a subclass of ionic liquids, a wealth of existing literature already describes the effect of structural modifications, to either of the ionic components, on the thermophysical properties of the resulting salts. This knowledge can be readily applied to this field to guide design with performance and broad sustainability considerations in mind. Probing the relationships between the increased degrees of freedom at the phase transitions, the ion structure, and the volume change should provide insights into the mechanisms responsible for the colossal barocaloric responses achieved here, allowing further advances of OIPCs in this context. With this broad scope of materials in view, we believe that OIPC-based materials have the potential to make substantial contributions to the sustainability of future cooling technologies.

## REFERENCES AND NOTES

1. International Energy Agency (IEA), "The Future of Cooling: Opportunities for energy-efficient air conditioning" (IEA, 2018).
2. Intergovernmental Panel on Climate Change (IPCC), Ed., *Climate Change 2022—Mitigation of Climate Change: Working Group III Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2023); https://www.cambridge.org/core/product/2929481A59B59C57C743A79420A2F9FF.
3. M. O. McLinden, J. S. Brown, R. Brignoli, A. F. Kazakov, P. A. Domanski, *Nat. Commun.* **8**, 14476 (2017).
4. International Energy Agency, "Sustainable, affordable cooling can save tens of thousands of lives each year" (IEA, 2023); https://www.iea.org/reports/sustainable-affordable-cooling-can-save-tens-of-thousands-of-lives-each-year.
5. X. Moya, N. D. Mathur, *Science* **370**, 797–803 (2020).
6. J. Liu, T. Gottschall, K. P. Skokov, J. D. Moore, O. Gutfleisch, *Nat. Mater.* **11**, 620–626 (2012).
7. X. Qian *et al.*, *Nature* **600**, 664–669 (2021).
8. L. Mañosa, A. Planes, *Adv. Mater.* **29**, 1603607 (2017).
9. K. Alex Müller *et al.*, *Appl. Phys. Lett.* **73**, 1056–1058 (1998).
10. B. Li *et al.*, *Nature* **567**, 506–510 (2019).
11. P. Lloveras *et al.*, *Nat. Commun.* **10**, 1803 (2019).
12. M. Romanini *et al.*, *Adv. Mater.* **33**, e2008076 (2021).
13. J. Seo *et al.*, *Nat. Commun.* **13**, 2536 (2022).
14. Y.-H. Gao *et al.*, *Nat. Commun.* **15**, 1838 (2024).
15. J. Seo *et al.*, *J. Am. Chem. Soc.* **146**, 2736–2747 (2024).
16. P. Lloveras, J.-L. Tamarit, *MRS Energy Sustain.* **8**, 3–15 (2021).
17. C. Aprea, A. Greco, A. Maiorino, C. Masselli, *Energy* **190**, 116404 (2020).
18. L. Cirillo, A. Greco, C. Masselli, *Therm. Sci. Eng. Prog.* **33**, 101380 (2022).
19. L. P. Rebelo, J. N. Canongia Lopes, J. M. Esperança, E. Filipe, *J. Phys. Chem. B* **109**, 6040–6043 (2005).
20. L. Jin *et al.*, *J. Am. Chem. Soc.* **134**, 9688–9697 (2012).
21. F. Chen, L. Jin, S. W. de Leeuw, J. M. Pringle, M. Forsyth, *J. Chem. Phys.* **138**, 244503 (2013).
22. A. Sourjah *et al.*, *Phys. Chem. Chem. Phys.* **25**, 16469–16482 (2023).
23. S. Das, A. Mondal, C. M. Reddy, *Chem. Soc. Rev.* **49**, 8878–8896 (2020).
24. H. Zhu, D. R. MacFarlane, J. M. Pringle, M. Forsyth, *Trends Chem.* **1**, 126–140 (2019).
25. D. R. MacFarlane, P. Meakin, N. Amini, M. Forsyth, *J. Phys. Condens. Matter* **13**, 8257–8267 (2001).
26. H. Ishida, N. Matsuhashi, R. Ikeda, D. Nakamura, *J. Chem. Soc., Faraday Trans. I* **85**, 111–120 (1989).
27. M. Lee *et al.*, *J. Mater. Chem.* **21**, 12280–12287 (2011).
28. U. A. Rana, R. Vijayaraghavan, D. R. MacFarlane, M. Forsyth, *Chem. Commun.* **47**, 6401–6403 (2011).
29. N. Sirigiri *et al.*, *Mater. Today Phys.* **22**, 100603 (2022).
30. M. Herstedt *et al.*, *J. Mol. Struct.* **783**, 145–156 (2006).
31. L. Jin *et al.*, *J. Phys. Chem. B* **121**, 5439–5446 (2017).
32. M. Zeng *et al.*, *Adv. Sci. (Weinh.)* **11**, e2306488 (2024).
33. D. Al-Masri, R. Yunis, A. F. Hollenkamp, C. M. Doherty, J. M. Pringle, *Phys. Chem. Chem. Phys.* **22**, 18102–18113 (2020).
34. Materials and methods are available as supplementary materials.
35. J. Li *et al.*, *Adv. Funct. Mater.* **31**, 2105154 (2021).
36. H. Zhu *et al.*, *J. Phys. Chem. Lett.* **9**, 3904–3909 (2018).
37. X. Lan *et al.*, *ACS Appl. Polym. Mater.* **4**, 9368–9377 (2022).
38. J. Huang, A. Hill, M. Forsyth, D. MacFarlane, A. Hollenkamp, *Solid State Ion.* **177**, 2569–2573 (2006).
39. A. Warrington *et al.*, *Mater. Chem. Front.* **6**, 1437–1455 (2022).
40. W. A. Henderson, V. G. Young Jr., S. Passerini, P. C. Trulove, H. C. De Long, *Chem. Mater.* **18**, 934–938 (2006).
41. W. A. Henderson *et al.*, *Inorg. Chem.* **45**, 1412–1414 (2006).
42. P. Lloveras, Ed., *Barocaloric Effects in the Solid State* (IOP Publishing, 2023), pp. 2053–2563.
43. P. G. Debenedetti, F. H. Stillinger, *Nature* **410**, 259–267 (2001).
44. A. Aznar *et al.*, *J. Mater. Chem. A Mater. Energy Sustain.* **8**, 639–647 (2020).
45. Y. Djikaev, A. Tabazadeh, P. Hamill, H. Reiss, *J. Phys. Chem. A* **106**, 10247–10253 (2002).
46. R. A. Shaw, A. J. Durant, Y. Mi, *J. Phys. Chem. B* **109**, 9865–9868 (2005).
47. B. Mason, *Adv. Phys.* **7**, 221–234 (1958).
48. I. Shamseddine, F. Pennec, P. Biwole, F. Fardoun, *Renew. Sustain. Energy Rev.* **158**, 112172 (2022).
49. M. H. Zahir, S. A. Mohamed, R. Saidur, F. A. Al-Sulaiman, *Appl. Energy* **240**, 793–817 (2019).
50. A. Fröba *et al.*, *Int. J. Thermophys.* **31**, 2059–2077 (2010).
51. J. M. P. França, C. A. Nieto de Castro, A. A. H. Pádua, *Phys. Chem. Chem. Phys.* **19**, 17075–17087 (2017).
52. Q. Ren *et al.*, *Nat. Commun.* **13**, 2293 (2022).
53. J. Seo, J. D. Braun, V. M. Dev, J. A. Mason, *J. Am. Chem. Soc.* **144**, 6493–6503 (2022).
54. S. Abeysooriya, F. Makhlooghiazad, J.-N. Chotard, L. A. O'Dell, J. M. Pringle, *J. Phys. Chem. C Nanomater. Interfaces* **127**, 12304–12320 (2023).

## ACKNOWLEDGMENTS

P. Newman is acknowledged for his technological support in developing the high-pressure differential thermal analyzer. C. Forsyth is acknowledged for his support in processing XRD data. The authors acknowledge use of facilities within the Monash X-ray Platform. **Funding:** This work was supported by the Australian Research Council (ARC) through Discovery Grant DP210101269. **Author contributions:** J.M.P., D.R.M., and M.K. conceptualized and acquired funding for the project. E.F.M., X.X., and K.-F.A.-Z. conducted preliminary high-pressure experiments. J.M.P., D.R.M., M.K., S.L.P., L.M., and A.S. contributed to method development. Investigation and manuscript composition were carried out by S.L.P., under the supervision of J.M.P., D.R.M., and M.K. All authors gave input into the final manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Crystallographic information files (CIFs) are available on the Cambridge Crystallographic Data Centre (CCDC) database; CCDC identifiers are noted in the supplementary materials. All other data are available in the manuscript or the supplementary materials. **License information:**

**SUPPLEMENTARY MATERIALS**

science.org/doi/10.1126/science.adq8396
Materials and Methods
Supplementary Text
Figs. S1 to S23
Tables S1 to S19
References (*55*–*68*)

Submitted 6 June 2024; accepted 29 October 2024
10.1126/science.adq8396

---

ELECTRONIC MATERIALS

# Surface conduction and reduced electrical resistivity in ultrathin noncrystalline NbP semimetal

**Asir Intisar Khan[1], Akash Ramdas[2]†, Emily Lindgren[2,3]†, Hyun-Mi Kim[4]†, Byoungjun Won[5], Xiangjin Wu[1], Krishna Saraswat[1], Ching-Tzu Chen[6], Yuri Suzuki[3,7], Felipe H. da Jornada[2,8], Il-Kwon Oh[5]*, Eric Pop[1,2,7,9]***

**The electrical resistivity of conventional metals such as copper is known to increase in thin films as a result of electron-surface scattering, thus limiting the performance of metals in nanoscale electronics. Here, we find an unusual reduction of resistivity with decreasing film thickness in niobium phosphide (NbP) semimetal deposited at relatively low temperatures of 400°C. In films thinner than 5 nanometers, the room temperature resistivity (~34 microhm centimeters for 1.5-nanometer-thick NbP) is up to six times lower than the resistivity of our bulk NbP films, and lower than conventional metals at similar thickness (typically about 100 microhm centimeters). The NbP films are not crystalline but display local nanocrystalline, short-range order within an amorphous matrix. Our analysis suggests that the lower effective resistivity is caused by conduction through surface channels, together with high surface carrier density and sufficiently good mobility as the film thickness is reduced. These results and the fundamental insights obtained here could enable ultrathin, low-resistivity wires for nanoelectronics beyond the limitations of conventional metals.**

Ultrathin conductors with low electrical resistance are needed for hyperscaled nanoelectronics (*1*), from metal interconnects for dense logic and memory (*2*, *3*) to neuromorphic (*4*) and spintronic devices (*5*, *6*). Low resistance allows lower voltage drops and lower signal delays, reducing power dissipation at the system level (*7*). Resistance is proportional to resistivity, but the resistivity of conventional metals increases in films or wires thinner than the electron mean free path (few tens of nanometers at room temperature) because of electron-surface scattering (*8*). For example, the room temperature resistivity of sub-5-nm thin Cu or Ru films is up to an order of magnitude larger than in bulk films (>100 nm) (*8*–*10*). High electrical resistivity of ultrathin metals can be a key contributor to energy consumption in dense logic and memory (*11*, *12*) and could ultimately limit the performance of future data-driven applications (*4*).

In this context, the topological Weyl semimetals NbP, NbAs, TaP, and TaAs (*13*–*18*) are promising because they could carry current within surface states that are topologically protected from disorder scattering (*19*). The multifold fermion semimetals CoSi, RhSi, AlPt, and GaPd have also been theoretically predicted (*20*, *21*) to benefit from surface conduction with suppressed scattering (*20*). In other words, as the thickness of such semimetals is reduced, the surface contribution to conduction (*22*) could lead to lower effective resistivity (*12*, *20*, *23*), whereas in conventional metals with nanoscale thickness, the electrons undergo more surface scattering (*8*, *11*, *24*). For example, recent measurements of high-quality crystalline NbAs displayed more than an order of magnitude reduction in the effective resistivity of nanobelts, reaching ~2 μohm·cm for ~250-nm thickness compared with their bulk single-crystal value of ~35 μohm·cm (*23*) at room temperature.

Surface-dominated transport has also been recently reported in amorphous $Bi_2Se_3$ topological insulator films (>75-nm thick) without long-range order (*25*), and disordered Weyl semimetal $WTe_x$ films (*26*) have shown good charge-to-spin conversion and electrical conductivity that are comparable to those of crystalline $WTe_2$ (*27*). Such experimental demonstrations with amorphous topological insulators suggest the possibility of surface-state conduction in Weyl semimetals even in the absence of long-range order. However, it is not known whether disordered or noncrystalline semimetals in ultrathin films (i.e., sub-5 nm) maintain surface-dominated transport and could be used to realize low-resistivity materials beyond the limitations of conventional metals. Such noncrystalline semimetals are much more likely to be compatible with modern semiconductor processing and ultradense future electronics, in which limited thermal budgets (<500°C) pose challenges for depositing single-crystal materials.

In this work, we uncovered a reduction of electrical resistivity in noncrystalline NbP semimetal with decreasing thickness down to ~1.5 nm. We found lower effective resistivity in sub-5-nm thin NbP films compared with their bulk crystalline counterparts, which we attribute to a proportionally higher conduction through a surface channel in the ultrathin films.

## Film growth and resistivity

The NbP films were sputtered on sapphire and other substrates at 400°C, a temperature compatible with back-end-of-line (BEOL) semiconductor fabrication (*28*). As shown in Fig. 1A, a seed layer of Nb was used to reduce the lattice mismatch between the substrate and the NbP films (*29*) and to promote local short-range order, i.e., nanocrystallinity. All samples were capped in situ with a ~3- to 4-nm thick silicon nitride layer to limit surface oxidation (see the supplementary materials and methods, "materials deposition" section; fig. S1; and table S1). We used high-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) to image the cross section of the NbP/Nb thin films, revealing local short-range order and nanocrystallinity within an amorphous matrix in the NbP layer across various thicknesses (~18 nm in Fig. 1, B and C, and figs. S2 and S3 and ~1.5 to 4.3 nm in fig. S4). Energy-dispersive spectroscopy (EDS) and x-ray photoelectron spectroscopy (XPS) analysis confirmed the stoichiometry and the uniform distribution of Nb and P within our sputtered NbP samples (fig. S5). STEM, EDS, and XPS characterization methods are detailed in the supplementary materials and methods, "materials characterization" section.

We measured the in-plane electrical resistivity of our NbP/Nb films and control Nb samples using standard Hall and eddy current–based contactless methods (*30*) (for details, see the

---

[1]Department of Electrical Engineering, Stanford University, Stanford, CA, USA. [2]Department of Materials Science and Engineering, Stanford University, Stanford, CA, USA. [3]Geballe Laboratory for Advanced Materials, Stanford University, Stanford, CA, USA. [4]Korea Electronics Technology Institute, Seongnam-si, Republic of Korea. [5]Department of Intelligence Semiconductor Engineering, Ajou University, Suwon, Republic of Korea. [6]IBM T.J. Watson Research Center, Yorktown Heights, NY, USA. [7]Department of Applied Physics, Stanford University, Stanford, CA, USA. [8]Stanford Institute for Materials & Energy Sciences, SLAC National Accelerator Laboratory, Menlo Park, CA, USA. [9]Precourt Institute for Energy, Stanford University, Stanford, CA, USA.
***Corresponding author. Email: epop@stanford.edu (E.P.); ikoh@ajou.ac.kr (I.-K.O.)**
†These authors contributed equally to this work.

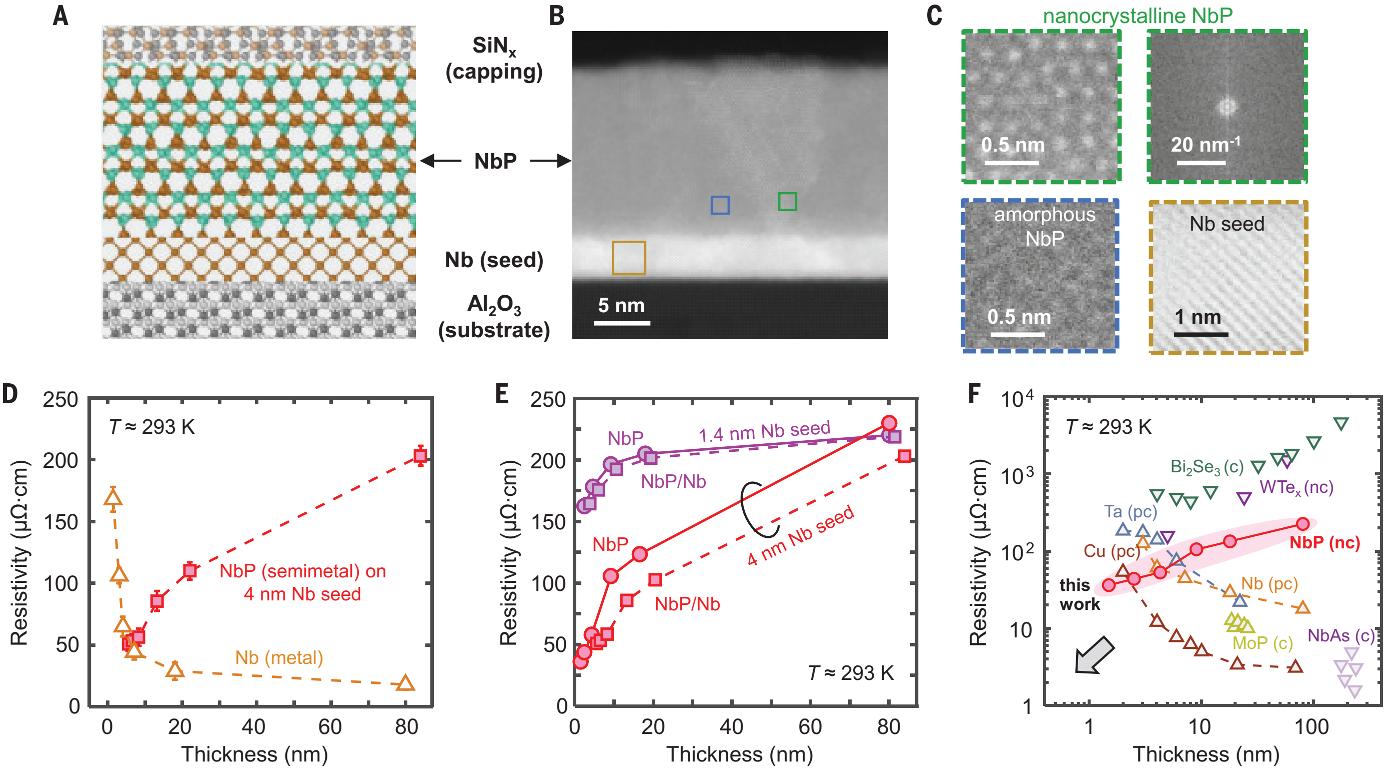


**Fig. 1. NbP/Nb thin film stacks and room temperature resistivity.** (**A**) Schematic of the sputtered NbP/Nb film stack. (**B**) HAADF-STEM cross section of an ~18-nm NbP/Nb film stack. The $Al_2O_3$ substrate, thin (~4-nm) Nb seed layer, and silicon nitride capping can also be seen. (**C**) Magnified STEM images show local short-range ordering and nanocrystallinity within the amorphous NbP layer, as well as the crystallinity of the Nb seed layer (see fig. S2 for other NbP thicknesses). (**D**) Room temperature resistivity versus thickness of NbP/Nb films (squares) and control Nb films (triangles). The resistivity and thickness of NbP plotted here includes the 4-nm Nb seed layer. The NbP/Nb stack showed unconventional resistivity scaling in that the effective resistivity decreased in thinner films. Symbols and error bars mark the average and SD, respectively, across five samples of each film thickness. (**E**) Room temperature resistivity versus thickness of NbP/Nb stacks before (squares) and after (circles) subtracting the Nb seed-layer conduction contribution. Samples with two different Nb seed layers are shown: 4 nm (red) and 1.4 nm (violet). Unconventional resistivity scaling is noted for all films, both including and excluding the Nb seed layer contribution. The horizontal axis represents either the total stack thickness (NbP + Nb) or just the NbP thickness. (**F**) Room temperature resistivity versus thickness for various materials. Here, our sputtered NbP semimetal resistivity is shown after subtracting the contribution of the Nb seed; similarly, Cu resistivity is shown without the contribution of its liner and barrier layers (*48*). Other films include Nb (from this work); Ta; the crystalline topological insulator $Bi_2Se_3$ (*32*); the topological semimetals: noncrystalline $WTe_x$ and crystalline NbAs (*23*, *26*); and the topological metal MoP (*45*). The arrow marks the best corner region of smallest resistivity at low film thickness. c, pc, and nc (in parentheses) refer to crystalline, polycrystalline, and noncrystalline films, respectively. The sputtered NbP displayed decreasing resistivity down to sub-5-nm thickness, with the lowest resistivity in ultrathin films.

supplementary materials and methods, "device fabrication and electrical measurement" section). The control Nb samples were prepared with the same deposition conditions as the Nb seed layers beneath the NbP samples. Figure 1D shows that the measured total room temperature resistivity of NbP/Nb films decreased from ~200 μohm·cm for ~80-nm thick NbP to ~51 μohm·cm for ~1.5-nm thick NbP (all on 4-nm Nb). This resistivity plot includes the electrical and thickness contribution of the 4-nm seed Nb layer. However, the resistivity of our control Nb metal films increased substantially as their thickness was reduced over the same range.

The measured temperature dependence of total resistivity shown in fig. S6A revealed metallic behavior (resistivity proportional to temperature) in NbP films of 18 nm or thinner, here including the 4-nm Nb seed layer (shown separately in fig. S6B). By contrast, an ~80-nm NbP film (also on a 4-nm Nb seed) showed resistivity that was almost independent of temperature, a signature of disorder or impurity-dominated bulk states (*25*). The reduced effective resistivity of the thinner NbP films suggested that there may be a nonnegligible contribution from surface carriers to the total conductance of these samples (*32*, *33*, *34*), which is explored in more detail in Fig. 3.

Figure 1E shows that the unconventional resistivity scaling with thickness in our NbP/Nb film was preserved for varying thicknesses of the Nb seed layer (4 and 1.4 nm). This decreasing resistivity with decreasing film thickness was also observed after the conductance of the thin Nb seed layer (fig. S6C) was subtracted from that of the NbP/Nb stack (fig. S6D), indicating that the NbP film was responsible for the observed trend seen in Fig. 1E. For comparison, we also prepared Cu/Nb films with similar thickness; fig. S7 shows that their resistivity increased as their thickness was reduced both before and after subtracting the conductance contribution of the 4-nm Nb. In other words, the Nb layer did not influence the contrasting resistivity trend observed for NbP versus Cu. Figure 1E also shows that the resistivity of NbP on the 1.4-nm Nb seed was higher than for NbP on the 4-nm Nb seed, which we attributed to the relatively lower

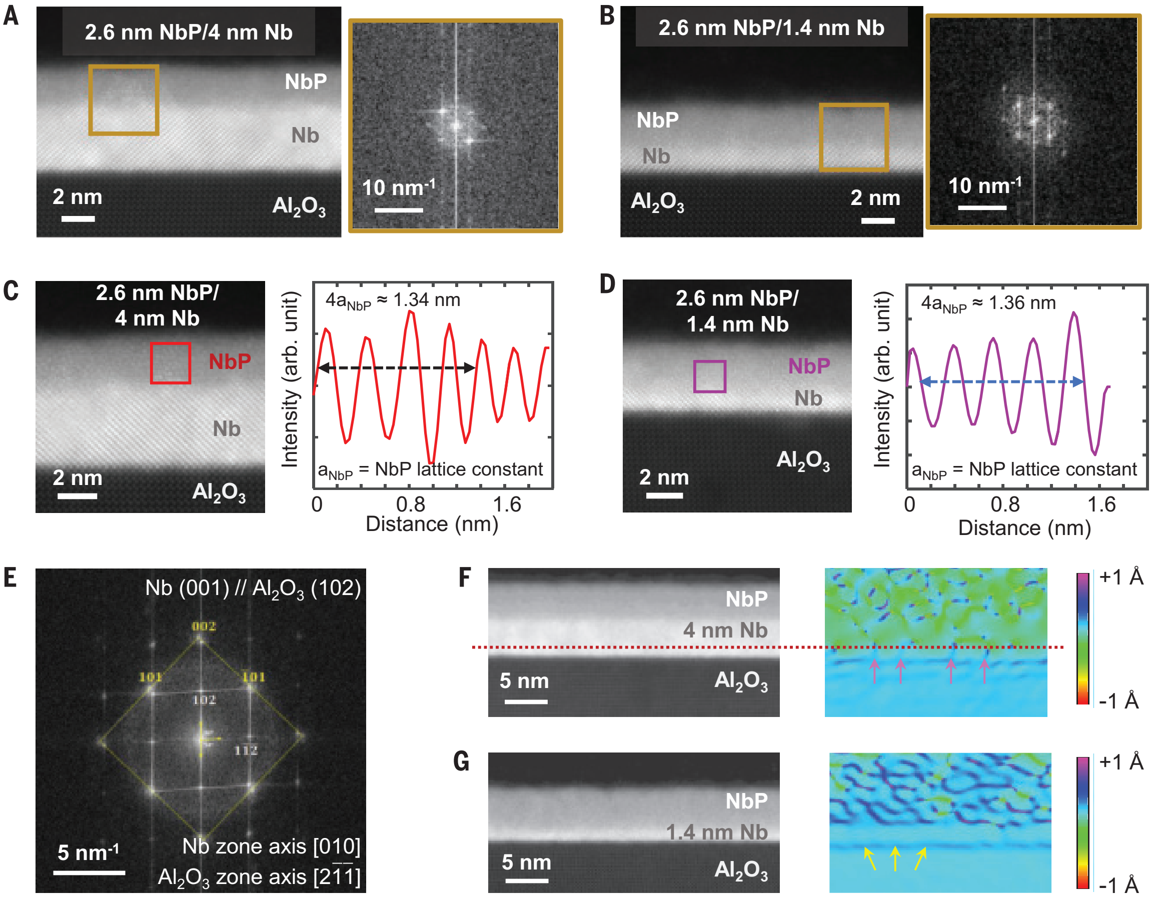


**Fig. 2. Microstructure details of ultrathin NbP/Nb heterostructures.** (**A** and **B**) HAADF-STEM images and their fast Fourier transform (FFT) diffraction of 2.6-nm NbP films on Nb seed layer with 4-nm thickness (A) and 1.4-nm thickness (B). (**C**) Local nanocrystalline (short-range order) region of 2.6-nm thin NbP film on 4-nm Nb seed, showing a NbP lattice constant of ~3.34 Å, near its nominal value of ~3.332 Å (*40*). (**D**) Similar image on a 1.4-nm Nb seed layer, revealing a NbP lattice constant of ~3.41 Å, which indicated that NbP was strained on the thinner Nb seed. (**E**) Diffraction pattern of Nb seed layer and $Al_2O_3$ substrate. Nb seed layers have an epitaxial relationship with the $Al_2O_3$ substrate – Nb (001) || $Al_2O_3$ (102). The Al in $Al_2O_3$ (102) has a rhombus lattice tilted by 6° compared with Nb (001). (**F** and **G**) Lattice strain analysis of 2.6-nm NbP film on 4-nm Nb (F) and on 1.4-nm Nb (G) rom Fourier filtering the corresponding HAADF-STEM images. The 1.4-nm Nb seed was strained laterally along the $Al_2O_3$ surface (yellow arrows), but the accumulated strain was released in the 4-nm Nb seed by forming misfit dislocations (pink arrows). Red dotted line marks the level of dislocations within the Nb seed. The colored images display the strain mapping of the layers. The greater green proportion in the top plot marks a larger unstrained portion of NbP on the thicker (~4-nm) Nb seed compared with the thinner (~1.4-nm) one.

strain in NbP with the thicker Nb seed, as discussed further below. The room temperature resistivity of our sub-3-nm thin NbP films on the 4-nm Nb seed was <45 μohm·cm, lower than the crystalline, bulk NbP resistivity of ~60 to 70 μohm·cm (*14*, *29*). The thinnest, 1.5-nm NbP film reached 34 μohm·cm at room temperature (red circles) after subtracting the Nb seed contribution. This is over six times lower than the bulk NbP resistivity in our thickest samples (Fig. 1E).

Figure 1F shows the scaling of room temperature resistivity versus thickness in our noncrystalline NbP semimetal, revealing a trend unlike traditional metals such as Cu, Nb, and Ta, and achieving one of the lowest resistivities at sub-5-nm thickness. We also quantify the total sheet resistance, $R_\square$, versus thickness of various films in fig. S8, including their seed or barrier layers, if any. As total thickness decreases from ~20 down to ~5 nm, the $R_\square$ of conventional metals increases by 10 to 100 times, but the $R_\square$ of topological semimetals increases by less than a factor of 2. Previously, resistivity smaller than the bulk resistivity was detected in NbAs nanobelts (*23*), topological insulators such as $Bi_2Se_3$ (*32*), and multifold fermion semimetal CoSi nanowires (*35*), although such films displayed greater crystallinity, greater thicknesses, and were deposited at higher temperature (typically >600°C). Multilayer graphene can also reach low resistivity in nanometer thin films, but only with substantial doping (*36*, *37*) and with high-temperature growth and processing (*38*). By contrast, the low deposition temperature (400°C) of our noncrystalline NbP films is compatible with industrial BEOL processes, a key advantage for integration into state-of-the-art nanoelectronics (*28*, *39*).

We also measured low resistivity and a similar resistivity scaling trend in ultrathin NbP films on different substrates such as MgO and $SiO_2$/Si (fig. S9A), as well as with different capping layers including silica and alumina. In terms of stability, uncapped NbP thin films (~2.6 nm) on 4-nm Nb measured in air showed a <10% change of resistivity after 4 days versus ~90% change in 4-nm Nb metal films, indicating less surface oxidation of NbP (fig. S9B). This stability is also promising for interconnect applications.

## Structural studies

The resistivity of the sub-20-nm NbP thin films on 4-nm Nb seeds was notably lower

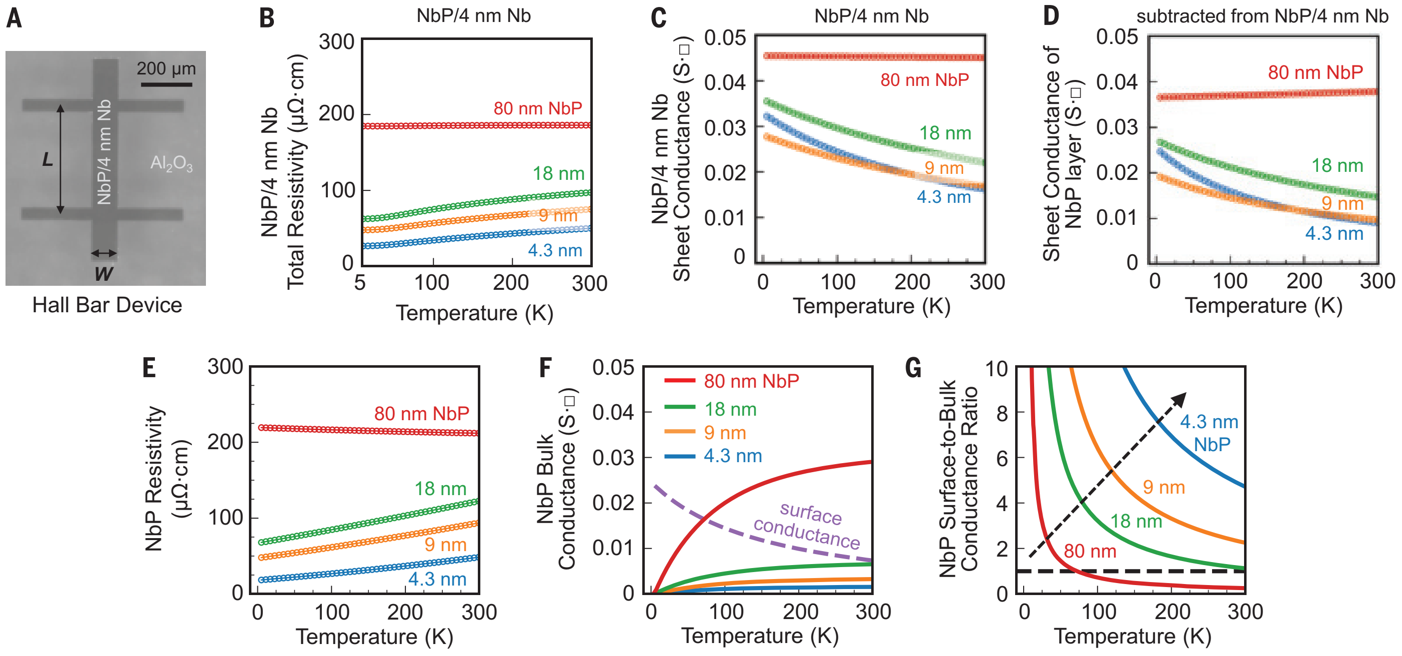


**Fig. 3. Temperature-dependent transport of NbP/Nb and NbP.** (**A**) Top view optical image of the Hall bar with width *W* of 100 μm and length *L* of 400 μm. The NbP was seeded by Nb and capped by $SiN_x$, as in Fig. 1B. (**B** and **C**) Temperature-dependent total resistivity of NbP/Nb (B) and sheet conductance of NbP/Nb films with varying NbP thicknesses (4.3, 9, 18, and 80 nm) on a 4-nm Nb seed (C). (**D**) Sheet conductance of the NbP layer of varying thicknesses obtained by subtracting the conductance of the 4-nm Nb seed layer (measured separately; see fig. S4A) from the total sheet conductance of NbP/Nb films in (C). (**E**) Temperature-dependent resistivity of NbP films with varying thicknesses, from 4.3 to 80 nm [obtained using (D)]. (**F**) Two-channel conductance fit to the resistivity data in (E), indicating a metallic surface channel conductance (dashed line) and disorder-dominated bulk conductance (solid lines). Here, we assumed the surface channel has zero thickness. Figure S13B displays the fit with a finite surface thickness of ~5 Å, yielding a similar result. (**G**) Calculated surface-to-bulk conductance ratio versus temperature for NbP films. The surface-to-bulk conductance ratio increased as the NbP film thickness was reduced (indicated by the dashed black arrow) across a wide range of temperatures. The region above the dashed line was dominated by surface conduction.

than that of NbP on the 1.4-nm Nb seed. To understand this difference, we imaged ~2.6-nm thin NbP films on Nb seed layers with 4- and 1.4-nm thicknesses (Fig. 2, A and B, respectively) using atomic-resolution HAADF-STEM. Magnified STEM images and corresponding diffraction patterns show the presence of similar nanocrystallinity within the amorphous matrices of NbP on both Nb seed layers (also see fig. S2 for 18-nm NbP films). Both 4- and 1.4-nm Nb seed layers were crystalline (fig. S3). NbP films were predominantly amorphous, with several nanometer-sized crystalline regions regardless of the Nb seed-layer thickness. Thus, the observed NbP resistivity scaling with thickness (Fig. 1E) for varying Nb seed layers would not likely be affected by the microstructure of the NbP films. The average lattice constant of our ~2.6-nm thin NbP film on the 4-nm Nb seed layer (Fig. 2C) was ~3.34 Å (~3.33 Å for ~18-nm NbP film; fig. S10A), near that of single-crystal NbP (*40*). However, Fig. 2D and fig. S10B show that the NbP film was strained, with higher average lattice constant (~3.41 Å for ~2.6-nm NbP and ~3.5 Å for ~18-nm NbP) on the 1.4-nm Nb seed layer, which could cause the higher resistivity (*41*, *42*) seen for ultrathin NbP on the 1.4-nm Nb seed layer (Fig. 1E).

We further found (Fig. 2E) that the epitaxial relationship between the Nb seed and the $Al_2O_3$ substrate was Nb (001) || $Al_2O_3$ (102). The Al in $Al_2O_3$ (102) had a rhombus lattice tilted by 6° compared with the square lattice of the Nb (100) plane. As a result, in-plane misfit strain occurred between the Nb seed and the substrate (fig. S11). Increasing the Nb seed layer thickness generated misfit dislocations within the Nb that released this strain energy. We observed strain release in the films with ~4-nm Nb seed (Fig. 2F), where the Nb lattice returned to its cubic structure with nominal lattice constant of ~3.32 Å. For the thinner 1.4-nm Nb seed, the misfit dislocations that could release stress were not observed (Fig. 2G). This laterally strained the 1.4-nm Nb seed layer with a lattice constant of ~3.53 Å, near that of the $Al_2O_3$ substrate; therefore, the NbP films on the 1.4-nm Nb seed also display lateral strain (Fig. 2D and fig. S10B), and the strained NbP/Nb interface could also cause charge scattering, further increasing the resistivity of the tensile NbP films (*41, 42*) on the 1.4-nm Nb seed (Fig. 1E).

### Transport measurements

As our next step, we wished to understand what causes the unusual resistivity scaling trend (versus thickness) in our NbP semimetal films. Previous reports had suggested surface-dominated conduction in topological insulators ($Bi_2Se_3$) and topological semimetals (TaAs, NbAs) in both their crystalline (*23*, *32*, *33*) and amorphous or nanocrystalline $Bi_2Se_3$ (*25*, *31*) films, attributed to topologically protected surface states. As the sample thickness decreases, conduction dominated by such surface states could explain the reduced resistivity of our thinner NbP films compared with their thicker counterparts. To understand this, we performed temperature-dependent transport measurements for a series of NbP thin films with varying thicknesses (~80 to ~4.3 nm) on the 4-nm Nb seed using standard Hall bar devices (Fig. 3A and see the supplementary materials and methods).

The unconventional trend of decreasing resistivity with decreasing NbP/Nb sample thickness persisted across all temperatures probed down to 5 K (Fig. 3B). The three thinner NbP/Nb films (4.3-, 9-, and 18-nm NbP, each on 4-nm Nb) showed decreasing resistivity with decreasing temperature (metallic behavior). By contrast,

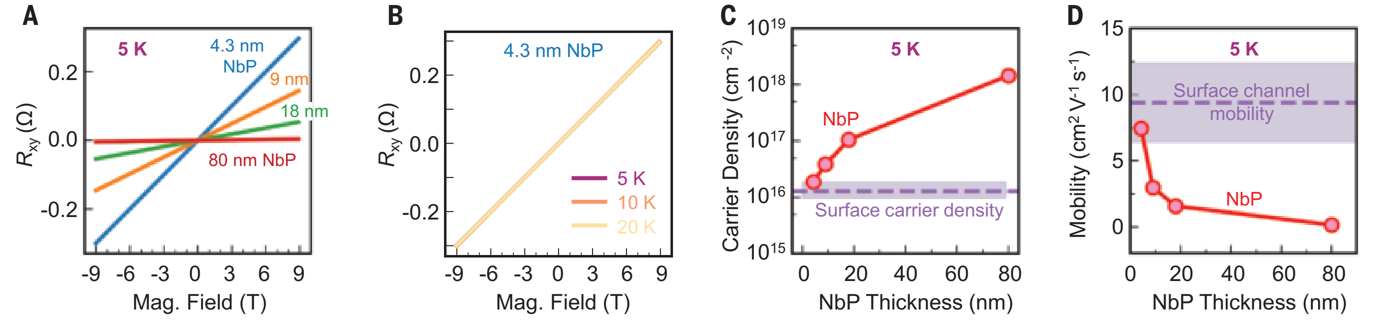


**Fig. 4. Hall measurements and carrier densities of our NbP films.** (**A**) Hall resistance versus magnetic field for NbP films with varying thicknesses at 5 K. (**B**) Hall resistance of a 4.3-nm thin NbP film versus magnetic field at 5, 10, and 20 K. (**C**) Sheet carrier density (holes, extracted from Fig. 4A) shows reduction with NbP film thickness. From the Hall coefficient versus thickness fit in fig. S17, we estimate a surface carrier density of 1.4 ± 0.4 × $10^{16}$ cm$^{-2}$, the sheet carrier density in the limit of zero NbP film thickness (shaded purple region). (**D**) Mobility of the NbP films showing an increasing trend with decreasing thicknesses. The shaded region represents the range of the surface channel mobility, 9.4 ± 3.0 cm$^2$ V$^{-1}$ s$^{-1}$, estimated from the surface carrier density. All data and estimates in this figure were obtained after subtracting the conduction contribution of the 4-nm Nb seed (see the supplementary materials and methods and fig. S15). Including the conduction contribution of the 4-nm Nb seed layer did not alter the carrier density and mobility trends shown in (C) and (D) (fig. S18).

the thick NbP/Nb film (~80-nm NbP on the 4-nm Nb seed) displayed a resistivity that was almost independent of temperature, a signature of disorder- or impurity-dominated bulk states (*25*). The reduced resistivity in the thinner NbP/Nb films that was maintained down to ~5 K suggested a non-negligible contribution of surface conduction in these samples (*32*–*34*).

To obtain the sheet conductance of the NbP layer (Fig. 3D), we subtracted the sheet conductance of the 4-nm Nb seed layer (fig. S6C) from the total sheet conductance of the NbP/Nb stack (Fig. 3C) over the 5 to 300 K temperature range. The extracted resistivity of the NbP layer also showed the unconventional trend of decreasing resistivity with decreasing NbP thicknesses from room temperature down to 5 K (Fig. 3E). To better understand the trend shown in Fig. 3E quantitatively, we fit the conductance of the NbP layers (with various thicknesses) with both bulk and surface channel contributions to the conductance between 5 and 300 K (Fig. 3F and figs. S12 and S13). We assumed that the NbP surface conductance contribution was constant with the sample thickness (for further details, see the supplementary text, "surface and bulk conductance of NbP/Nb and NbP layer" section, and fig. S12).

As can be seen from Fig. 3F and fig. S13, the bulk conductance of our NbP films increased from 5 to 300 K, as would be expected for variable-range hopping behavior in amorphous and nanocrystalline films (*25*). Conversely, the surface conductance was metallic and decreased with increasing temperature (*25*, *32*). As thickness decreased from ~80 to ~4.3 nm, the bulk contribution to the conductance decreased. At low temperatures, we expected the hopping carrier transport to be small and nearly independent of sample thickness. Thus, the conduction was dominated by a surface channel at low temperature (e.g., <50 K) even in the thicker 80-nm NbP sample (*25*, *32*).

We also estimated the surface-to-bulk conductance ratio (Fig. 3G), which revealed that all of the thinner films (18-nm NbP or less) were dominated by their surface contribution up to room temperature. The resistivity of our 4.3-nm NbP film was smaller than the bulk single-crystal NbP resistivity (*14*, *29*), whereas the resistivity of our 80-nm NbP film (Fig. 1E) was ~3× higher than the single-crystal value. The lower resistivity of our thinner NbP was unlikely to have been the result of improved crystallinity because these films were predominantly amorphous with embedded nanocrystallites (Fig. 2, A and B).

We also estimated the bulk NbP conductance and effective surface conductance of NbP (with the Nb layer) from the total sheet conductance of the NbP/Nb samples versus NbP thickness in fig. S14, with the analysis detailed in the supplementary materials and methods. Figure S14A shows that the surface conductance dominated the total sheet conductance for all NbP/Nb film stacks thinner than ~30 nm at room temperature. Even in the presence of defects or disorder, the higher conductivity in our thinner NbP/Nb films and NbP layers came from a surface-like channel.

### Carrier density estimates

We performed Hall resistance measurements of our NbP films as a function of magnetic field at 5 K (Fig. 4A). We subtracted the deduced Hall conductivity of the 4-nm Nb seed layer (fig. S15A) from that obtained for our stacks (fig. S15B). As shown in Fig. 4A, the Hall resistance was linear with magnetic field at all sample thicknesses, suggesting that a single carrier dominated transport in our NbP films (in this case, holes). The Hall resistance of our 4.3-nm thick NbP versus magnetic field was nearly independent of temperature between 5 and 20 K (Fig. 4B). The extracted sheet carrier density at 5 K shown in Fig. 4C decreased from ~$10^{18}$ cm$^{-2}$ for 80-nm thick NbP to ~$10^{16}$ cm$^{-2}$ in 4.3-nm thin NbP (for details, see the supplementary text). This trend was consistent with previous reports on thicker films of the crystalline topological semimetals NbAs and TaAs (*23*, *33*).

The carrier density per unit volume in our NbP films (>$10^{22}$ cm$^{-3}$; fig. S16) was higher (*43*) than that in NbP bulk single crystals (*14*) but comparable to other topological semimetals such as ~70-nm thick NbP epitaxial films (>$10^{22}$ cm$^{-3}$) (*29*), textured and amorphous CoSi (*44*), and topological metals such as MoP (>$10^{23}$ cm$^{-3}$) (*45*). In addition, the effective carrier density estimated from Hall measurements in disordered or noncrystalline films, such as our NbP, could be overestimated (and the mobility underestimated) due to possible contribution from hopping-like transport (*46*). This has been reported in organic semiconductors (*46*) and the topological insulator $Bi_2Se_3$, where the total carrier density estimated in noncrystalline films was ~10 times higher (*25*) than in its crystalline counterpart (*32*).

The carrier density versus thickness trend (Fig. 4C) allowed us to estimate an average surface carrier density of ~$10^{16}$ cm$^{-2}$, i.e., the hole density in the limit of the NbP film thickness approaching zero. This projected surface carrier density in our noncrystalline NbP was ~3 times larger than what was estimated in crystalline NbAs (*23*); however, it is consistent with the possibility of a higher apparent carrier density from Hall measurements in a noncrystalline system, as explained above.

The estimated mobility at 5 K (Fig. 4D) showed an increasing trend with decreasing NbP thickness. The effective mobility (at 5 K) of a 4.3-nm thin NbP film was ~7.4 $cm^2\ V^{-1}\ s^{-1}$, ~50 times greater than that of the 80-nm thick NbP film (~0.15 $cm^2\ V^{-1}\ s^{-1}$). Using the extrapolated surface sheet carrier density (Fig. 4C) and surface conductance (Fig. 3F), we estimated the mobility (see the supplementary text) of the surface-like channel to be 9.4 ± 3.0 $cm^2\ V^{-1}\ s^{-1}$. This higher surface mobility appears to enable the lower resistivity in our thinnest NbP films (Fig. 3E), where conduction is dominated by surface rather than bulk channels (Fig. 3G). These estimates were performed after careful subtraction of the 4-nm Nb seed layer contribution (fig. S6); however, we found that the thickness-dependent carrier density and mobility trends shown in Fig. 4, C and D, were maintained even when the Nb layer is included, i.e., in NbP/Nb heterostructures (fig. S18).

What are the origins of the surface-like conduction in these ultrathin noncrystalline films? This remains a partly open question, but we suggest a few possible causes. One possibility is the formation of disorder-tolerant Fermi arc–like surface states (*23*) even in noncrystalline topological materials (*43*, *47*). Another cause may be the existence of an interfacial free-electron gas-like state (*32*) near the NbP/Nb interface, where we observed local short-range ordering (Fig. 2, A and B, and fig. S4). For example, topological surface states are expected to be metallic-like in nature (*25*) and less sensitive to disorder scattering (*19*, *23*). The estimated surface mobility (~9.4 $cm^2\ V^{-1}\ s^{-1}$ at 5 K) of our noncrystalline NbP films was much lower than that of crystalline NbP (~$10^6$ $cm^2\ V^{-1}\ s^{-1}$ at ~2 K) (*14*) and topological insulators such as $Bi_2Se_3$ (~$10^3$ $cm^2\ V^{-1}\ s^{-1}$ at 1.5 K) (*32*). However, the surface mobility in our films is comparable to mobilities found in sub-10-nm thin polycrystalline $Bi_2Se_3$ (<10 $cm^2\ V^{-1}\ s^{-1}$ at 1.5 K) (*32*, *34*) and in thick amorphous $Bi_2Se_3$ (<20 $cm^2\ V^{-1}\ s^{-1}$ at 2 K) (*25*) with topological surface states. The low resistivity of our ultrathin NbP films was caused by the combination of high surface carrier density (~$10^{16}$ $cm^{-2}$) and sufficiently good surface mobility. The low effective resistivity that we found was surface dominated and maintained up to room temperature in all sub-18-nm thin films (Fig. 3). Looking ahead, we expect our work to motivate future efforts into imaging surface-state dispersion in amorphous or noncrystalline semimetals, for example, by using surface-sensitive techniques such as angle-resolved photoemission spectroscopy (ARPES) and spin-resolved ARPES (*25*).

In conclusion, we found that the resistivity of noncrystalline NbP films decreased substantially as the film thickness was reduced, which is a trend counter to that observed in most common metals. The thinnest films (<5 nm) displayed resistivities lower than conventional metals of similar thickness at room temperature. Measurements and modeling indicated that our NbP films thinner than ~18 nm were dominated by surface conduction up to room temperature, which is the origin of the effective resistivity decrease in thinner films. These films were deposited by large-area sputtering at relatively low temperatures (400°C) compatible with modern microelectronics processing. These results and the fundamental insights obtained here could enable ultrathin topological semimetals as low-resistivity interconnects in future high-density electronics.

## REFERENCES AND NOTES

1. S. Salahuddin, K. Ni, S. Datta, *Nat. Electron.* **1**, 442–450 (2018).
2. A. A. Vyas, C. Zhou, C. Y. Yang, *IEEE Trans. Nanotechnol.* **17**, 4–10 (2018).
3. M. M. Shulaker *et al*., *Nature* **547**, 74–78 (2017).
4. D. Ielmini, H.-S. P. Wong, *Nat. Electron.* **1**, 333–343 (2018).
5. S. Jung *et al*., *Nature* **601**, 211–216 (2022).
6. L. Zhu *et al*., *Adv. Funct. Mater.* **29**, 1805822 (2019).
7. V. Huang, D. Shim, H. Simka, A. Naeemi, in *2020 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2020), pp. 32.6.1–32.6.4; https://doi.org/10.1109/IEDM13553.2020.9371945.
8. D. Gall, *J. Appl. Phys.* **127**, 050901 (2020).
9. E. V. Barnat, D. Nagakura, P.-I. Wang, T.-M. Lu, *J. Appl. Phys.* **91**, 1667–1672 (2002).
10. E. C. Ko, J. Y. Kim, H. Rhee, K. M. Kim, J. H. Han, *Mater. Sci. Semicond. Process.* **156**, 107258 (2023).
11. D. Gall, in *2020 International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA)* (IEEE, 2020), pp. 112–113; https://doi.org/10.1109/VLSI-TSA48913.2020.9203700.
12. C.-T. Chen *et al*., in *2020 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2020), pp. 32.4.1–32.4.4; https://doi.org/10.1109/IEDM13553.2020.9371996.
13. N. P. Armitage, E. J. Mele, A. Vishwanath, *Rev. Mod. Phys.* **90**, 015001 (2018).
14. C. Shekhar *et al*., *Nat. Phys.* **11**, 645–649 (2015).
15. S.-Y. Xu *et al*., *Nat. Phys.* **11**, 748–754 (2015).
16. L. Rocchino *et al*., *Nat. Commun.* **15**, 710 (2024).
17. S.-Y. Xu *et al*., *Science* **349**, 613–617 (2015).
18. J. N. Nelson *et al*., *Matter* **6**, 2886–2899 (2023).
19. G. Resta, S.-T. Pi, X. Wan, S. Y. Savrasov, *Phys. Rev. B* **97**, 085142 (2018).
20. S.-W. Lien *et al*., *NPJ Quantum Mater.* **8**, 3 (2023).
21. N. B. M. Schröter *et al*., *Nat. Phys.* **15**, 759–765 (2019).
22. M. Breitkreiz, P. W. Brouwer, *Phys. Rev. Lett.* **123**, 066804 (2019).
23. C. Zhang *et al*., *Nat. Mater.* **18**, 482–488 (2019).
24. T. Zhou, D. Gall, *Phys. Rev. B* **97**, 165406 (2018).
25. P. Corbae *et al*., *Nat. Mater.* **22**, 200–206 (2023).
26. X. Li *et al*., *Matter* **4**, 1639–1653 (2021).
27. M. J. Mleczko *et al*., *ACS Nano* **10**, 7507–7514 (2016).
28. J. Schmitz, *Surf. Coat. Tech.* **343**, 83–88 (2018).
29. A. Bedoya-Pinto *et al*., *ACS Nano* **14**, 4405–4413 (2020).
30. J. Krupka, D. Nguyen, J. Mazierska, *Meas. Sci. Technol.* **22**, 085703 (2011).
31. L. He *et al*., *Nano Lett.* **12**, 1486–1490 (2012).
32. N. Bansal, Y. S. Kim, M. Brahlek, E. Edrey, S. Oh, *Phys. Rev. Lett.* **109**, 116804 (2012).
33. N. L. Nair *et al*., *Phys. Rev. B* **102**, 075402 (2020).
34. A. A. Taskin, S. Sasaki, K. Segawa, Y. Ando, *Phys. Rev. Lett.* **109**, 066803 (2012).
35. C.-I. Tsai *et al*., *Cryst. Growth Des.* **9**, 4514–4518 (2009).
36. J.-Z. Huang *et al*., *ACS Appl. Nano Mater.* **6**, 10680–10686 (2023).
37. J. Jiang, J. H. Chu, K. Banerjee, in *2018 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2018), pp. 34.5.1–34.5.4; https://doi.org/10.1109/IEDM.2018.8614535.
38. S. Vaziri *et al*., *IEEE Electron Device Lett.* **41**, 1592–1595 (2020).
39. C. Fenouillet-Beranger *et al*., *IEEE Trans. Electron Dev.* **68**, 3142–3148 (2021).
40. J. Xu *et al*., *Inorg. Chem.* **35**, 845–849 (1996).
41. S. C. Hunter, F. R. N. Nabarro, R. E. Peierls, *Proc. R. Soc. London. Ser. A.* **220**, 542–561 (1953).
42. K. Tanaka, T. Watanabe, *Jpn. J. Appl. Phys.* **11**, 1429 (1972).
43. T. Peng *et al*., *Phys. Rev. B* **106**, 125310 (2022).
44. A. Molinari *et al*., *ACS Appl. Electron. Mater.* **5**, 2624–2637 (2023).
45. H. J. Han *et al*., *Adv. Mater.* **35**, e2208965 (2023).
46. H. T. Yi, Y. N. Gartstein, V. Podzorov, *Sci. Rep.* **6**, 23650 (2016).
47. C. Wang, T. Cheng, Z. Liu, F. Liu, H. Huang, *Phys. Rev. Lett.* **128**, 056401 (2022).
48. A. Pyzyna *et al*., in *2017 IEEE International Interconnect Technology Conference (IITC)* (IEEE, 2017), pp. T120–T121; https://doi.org/10.1109/VLSIT.2015.7223712.

## ACKNOWLEDGMENTS

We dedicate this work in memory of the late Prof. A. K. M. Newaz (San Francisco State University) and Prof. E. Reed (Stanford University). We are grateful for discussions with them throughout the years and for their contribution to the scientific community. We also acknowledge Prof. H.-S. P. Wong for discussions and encouragement related to this work. A.I.K. is thankful to J. McVittie and C. Kline for their support and discussion regarding materials deposition, to S. Komera for the lab support, and to C. Lavoie for additional insights. The TEM work was supported by T. Cheon at the Daegu Gyeongbuk Institute of Science and Technology (DGIST). A.I.K. also thanks M. Noshin and H. Kwon for useful discussions on materials deposition. **Funding:** This work was performed at the Stanford Nanofabrication Facility (SNF) and Stanford Nano Shared Facilities (SNSF), which are supported by the National Science Foundation (NSF) award ECCS-2026822. The Stanford authors were supported in part by the Precourt Institute for Energy and the SystemX Alliance. A.I.K. acknowledges support from a Stanford Graduate Fellowship. E.L. and Y.S. were funded by the National Science Foundation (under grant no. 2037652). A.R. and F.H.J. acknowledge support from the NSF program Designing Materials to Revolutionize and Engineer our Future (DMREF project DMR-1922312). I.-K.O. acknowledges support from the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT grant RS-2024-00357895). **Author contributions:** A.I.K. conceived the idea with C.T.C., supported by E.P. and K.S. A.I.K. formulated and optimized the material deposition process with input from C.T.C. and K.S. A.I.K. led the design of experiments with input from E.P. A.I.K. performed the materials deposition with help from X.W. and input from E.P. TEM, EDS, XPS, and strain analysis were performed by H.-M.K., I.-K.O., and B.W. with lead and input from I.-K.O. A.I.K. fabricated the devices for transport measurements. E.L. performed the Hall measurements with input from Y.S., A.I.K., and A.R. The transport data were analyzed by A.R., A.I.K., F.H.J., and E.P. A.I.K., A.R., I.-K.O., and E.P. wrote the manuscript with input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in this paper are present in the main text or the supplementary materials.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adq7096
Materials and Methods
Supplementary Text
Figs. S1 to S18
Table S1
References

Submitted 27 May 2024; resubmitted 25 September 2024
Accepted 22 November 2024
10.1126/science.adq7096

GLOBAL WARMING

# Recent global temperature surge intensified by record-low planetary albedo

Helge F. Goessling[1]*, Thomas Rackow[2], Thomas Jung[1,3]

**In 2023, the global mean temperature soared to almost 1.5 kelvin above the preindustrial level, surpassing the previous record by about 0.17 kelvin. Previous best-guess estimates of known drivers, including anthropogenic warming and the El Niño onset, fall short by about 0.2 kelvin in explaining the temperature rise. Using satellite and reanalysis data, we identified a record-low planetary albedo as the primary factor bridging this gap. The decline is apparently caused largely by a reduced low-cloud cover in the northern mid-latitudes and tropics, in continuation of a multiannual trend. Further exploring the low-cloud trend and understanding how much of it is due to internal variability, reduced aerosol concentrations, or a possibly emerging low-cloud feedback will be crucial for assessing the present and expected future warming.**

After the last major El Niño event in 2015–2016, global mean warming was offset by a transition to persistent La Niña conditions (*1*). Since March 2023, however, global sea surface temperatures have broken records (*2*), well ahead of substantial contributions from the more moderate 2023–2024 El Niño (*3*). With the annual global mean surface temperature (GMST) close to 1.5 K above the preindustrial level, the North Atlantic, in particular, made headlines, with the average surface temperatures exceeding previous records by clear margins (*2*). Concurrently, in 2022–2023 the Antarctic sea ice extent, after decades of surprising stability (*4*), fell far below previous levels (*5*, *6*).

In addition to the onset of El Niño and the expected long-term warming due to anthropogenic greenhouse gases, several factors may have contributed to the anomalous GMST in 2023 (*3*): The 11-year solar cycle is approaching its intensity maximum (*7*); the submarine volcano Hunga Tonga–Hunga Ha'apai has released large amounts of water vapor into the stratosphere, although a net warming effect has been questioned (*8*); and new ship fuel regulations, aimed at reducing sulfur emissions, were implemented in three phases, in 2010, 2015, and 2020 (*9*). Although these regulations may be associated with a spatial pattern that is roughly consistent with the pronounced warming of the traffic-heavy North Atlantic and despite further evidence for recent warming due to reduced aerosols (*9*–*12*), it has been estimated that the combined global effect of all three factors is below 0.1 K (*3*). This contradicts claims of a much stronger shipping-related effect (*13*) that neglected the delayed response to forcing. An unexplained warming of about 0.2 K thus remains (*3*). Based on Clouds and the Earth's Radiant Energy System–Energy Balanced and Filled (CERES-EBAF; hereafter CERES) data [(*14*–*16*); materials and methods], the recent warming has been linked to an unusually large total top-of-atmosphere (TOA) Earth's energy imbalance [EEI; (*5*)].

We used CERES satellite and ERA5 [European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5] reanalysis [(*17*, *18*); materials and methods] data to explore the causes of the temperature surge. As synthesized in Fig. 1 and detailed below, we found that the unusually large recent imbalance was mainly driven by a record-low planetary albedo in 2023, continuing a multiannual trend related to decreasing shortwave reflection by clouds [consistent with (*19*, *20*)]. The cloud-related albedo reduction is apparently largely due to a pronounced decline of low-level clouds over the northern mid-latitude and tropical oceans, particularly the Atlantic. The increased absorption of shortwave radiation since December 2020 due to reduced albedo can explain 0.22 ± 0.04 K of the 2023 temperature anomaly, including 0.030 ± 0.006 K from polar regions, where declining albedo is dominated by the retreat of both sea ice and snow. Increased incident solar radiation associated with a strong solar-cycle maximum, captured by CERES but absent in ERA5, has contributed 0.027 ± 0.005 K to the 2023 temperature anomaly, whereas El Niño has added 0.07 ± 0.04 K. Disentangling contributions to the low-cloud trend from internal variability, indirect aerosol effects, and a possibly emerging low-cloud feedback remains challenging.

## Record-high EEI and planetary albedo

EEI has been positive for many decades owing to increasing greenhouse gas concentrations (*14*). According to CERES, the imbalance has been increasing since 2000 [(*5*, *9*, *19*, *21*); Fig. 2C], reaching a rate of +0.76 W m$^{-2}$ decade$^{-1}$ during the decade prior to 2023 (Table 1). However, a record high was reached in 2023 with an anomaly of +0.97 W m$^{-2}$ relative to 2001–2022. ERA5 agrees on the positive sign but exhibits lower values, particularly for the imbalance in 2023 (+0.31 W m$^{-2}$; Table 1 and Fig. 2C; see materials and methods for a discussion of the uncertainties inherent to ERA5 energy budget diagnostics).

The EEI trend and 2023 peak are not associated with decreasing outgoing longwave radiation (OLR), as one would expect from increasing greenhouse-gas concentrations in the absence of shortwave feedbacks. Instead, OLR has been increasing and largely offsetting even stronger absorbed solar radiation (ASR) anomalies [(*19*, *20*); Fig. 2D], consistent with climate models (*22*). The decadal 2013–2022 trend in ASR amounts to +1.10 W m$^{-2}$ decade$^{-1}$ in CERES and +0.97 W m$^{-2}$ decade$^{-1}$ in ERA5, followed by astonishing anomalies of +1.82 W m$^{-2}$ in CERES and +1.31 W m$^{-2}$ in ERA5 in 2023 (Table 1 and Fig. 2C). Variations of incident solar radiation (ISR), including by the 11-year solar cycle, are an order of magnitude smaller (*3*, *7*), which implies that reduced planetary albedo is the dominant cause (Table 1, Fig. 1A, and fig. S2B). However, it is notable that, according to CERES, ISR attained a positive anomaly in 2023 of +0.28 W m$^{-2}$, well above the previous solar-cycle maximum, whereas ERA5 forcing still assumed a negative anomaly of −0.08 W m$^{-2}$ (Table 1 and fig. S2F). Given an absolute planetary albedo of about 29% (fig. S1I), about 0.20 W m$^{-2}$ of the 2023 ASR and imbalance anomalies in CERES, and close to half of the discrepancies between CERES and ERA5, can be explained by the ISR peak.

Long-term trends in ERA5 can be spurious also owing to observing-system changes (*17*, *18*). ERA5, however, suggests that planetary albedo was possibly relatively low around the 1940s and '50s (fig. S3I), before industrial aerosol precursor emissions led to global dimming until the 1980s (*23*). The strongest planetary albedo excursions were high-albedo episodes caused by volcanic eruptions, with annual ASR anomalies reaching −3 W m$^{-2}$ in 1992 after the Mount Pinatubo eruption (fig. S3D). Negative albedo anomalies below the 1950s' minimum were, however, absent. Although uncertain, this suggests that the 2023 planetary albedo may have been the lowest since at least 1940.

## The most pronounced changes occurred in the Northern Hemisphere and tropics

Positive ASR anomalies in 2023 were most pronounced in the Northern Hemisphere and tropics (Fig. 3B and fig. S4A), consistent with the warming pattern (Fig. 3A). Regional maxima of the 2023 ASR anomaly, locally around 10 W m$^{-2}$, occurred over the eastern Indian Ocean, South America, and the eastern Pacific in the northern branch of the intertropical convergence zone, as well as over northern North

---

[1] Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany. [2] European Centre for Medium-Range Weather Forecasts, Bonn, Germany. [3] Institute of Environmental Physics, University of Bremen, Bremen, Germany.
*Corresponding author. Email: helge.goessling@awi.de

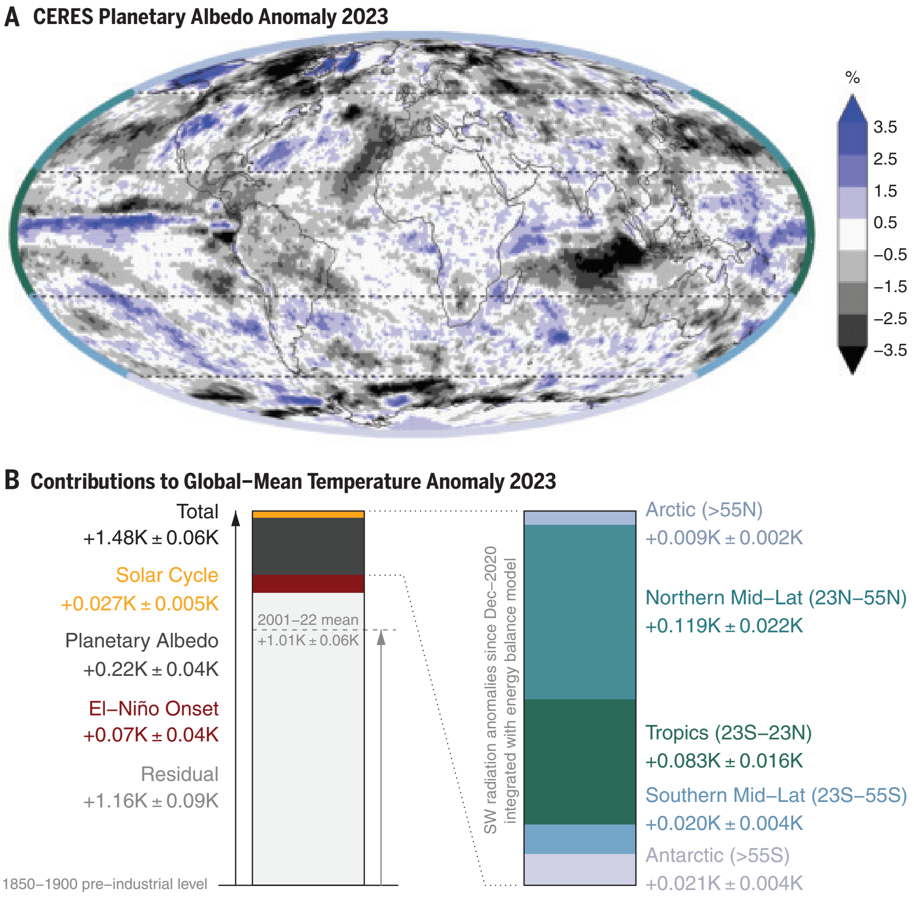
America, the Southern Ocean around 60°S, the subtropical and eastern North Atlantic, and parts of the North Pacific (Fig. 4D and fig. S6H). All of these anomalies are present in both datasets, but those over the North Atlantic and North Pacific are more pronounced in CERES, which may be related to the handling of aerosols in ERA5 (see the next sections).

The positive ASR anomalies over the northern extratropical oceans in 2023 are broadly consistent with decadal trends prior to 2023 (Fig. 4, C and D). This is not the case over the tropical Pacific and Indian Ocean, where interannual ASR anomalies are dominated by total cloud cover (TCC) changes associated with the El Niño–Southern Oscillation [ENSO; (*19*, *24*–*27*)], which transitioned from persistent La Niña to El Niño conditions after 2022 (Fig. 2B). A composite of nine El Niño events based on ERA5 data [materials and methods; Niño 3.4 index based on (*28*)] suggests that the albedo signature of El Niño onset years may contribute roughly 5% [+0.08 ± 0.06 W m$^{-2}$ relative to 2001–2022; fig. S7C] to the total 2023 ASR anomaly. Regionally, ENSO-related cloud and associated ASR patterns (fig. S7, I and J) largely explain inconsistencies between 2023 anomalies and 2013–2022 trends in the tropics (Fig. 4, C and D). This includes the strongly positive ASR anomalies (negative cloud anomalies) in the eastern Indian Ocean in 2023.

Before exploring cloud changes more generally, we considered the influence of surface albedo, which has been declining since the 1970s (fig. S3J), first primarily as a result of Arctic sea ice and snow retreat (*29*) and, since 2016, as a result of Antarctic sea ice retreat (*6*, *30*). This led to a pronounced seasonal signature in global mean surface albedo anomalies (fig. S2C) and polar ASR anomalies (Fig. 3B). In austral summer 2022–2023, the surface albedo anomaly of −0.4% was about as strong as the planetary albedo anomaly (fig. S2B). However, surface albedo anomalies are attenuated by about a factor of three on average, primarily because of cloud masking (*31*, *32*), and even more in the cloudy polar regions (*33*). Surface albedo thus contributed only weakly to the recent planetary albedo decline, particularly when averaged annually and globally, which is further quantified in the next sections.

## ASR anomalies are closely linked with cloud changes

Given the central role of clouds in Earth's radiation budget (*19*, *21*, *24*, *31*, *32*, *34*, *35*), spatial patterns of positive ASR anomalies and


**A** CERES Planetary Albedo Anomaly 2023

**B** Contributions to Global–Mean Temperature Anomaly 2023

**Fig. 1. Synthesis of contributions to Earth's surface temperature anomaly 2023.** (**A**) CERES planetary albedo anomaly 2023 relative to 2001–2022 (derived from annual mean TOA incident solar and upwelling solar radiation). (**B**) Contributions to the global mean surface temperature anomaly 2023 found in this study, with 90% confidence intervals (materials and methods). Shown on the left are contributions from the El Niño onset (red) and from ASR anomalies since December 2020 as a result of incident solar radiation anomalies (orange) and planetary albedo anomalies (dark gray). Shown on the right are the ASR-related contributions further decomposed into contributions from five zonal bands. SW, shortwave.

**Table 1. Decadal trends for 2013–2022 and anomalies for 2023 of selected global mean quantities related to Earth's temperature, clouds, and energy budget.** Trends are per decade, and 2023 anomalies are relative to 2001–2022. EEI represents Earth's TOA total energy imbalance, ASR represents the TOA net solar radiation, and CREtc represents the TOA solar CRE effect inferred from TCC anomalies. Planetary albedo is derived from global mean TOA solar downwelling and upwelling radiation. Bold numbers correspond to counterfactuals based on a two-layer EBM in which ASR anomalies are assumed to be zero from the beginning of December 2020 onward (as in Fig. 1 and fig. S2). Dashes indicate data not available.

| Data type | Surface (skin) temperature (K) | EEI (W m$^{-2}$) | ASR (W m$^{-2}$) | CREtc (W m$^{-2}$) | TCC fraction (%) | Low-cloud cover fraction (%) | Planetary albedo (%) | TOA ISR (W m$^{-2}$) |
|---|---|---|---|---|---|---|---|---|
| CERES trend 2013–2022 (per decade) | – | +0.76 | +1.10 | +0.47 | −0.37 | – | −0.33 | −0.01 |
| ERA5 trend 2013–2022 (per decade) | +0.24 | +0.18 | +0.97 | +1.24 | −1.16 | −1.27 | −0.35 | −0.34 |
| CERES anomaly 2023 | – (**+0.22**)* | +0.97 (**−0.44**) | +1.82 (**+0**) | +0.58 | −0.34 | – | −0.48 | +0.28 |
| ERA5 anomaly 2023 | +0.47 (**+0.31**) | +0.31 (**−0.74**) | +1.33 (**+0**) | +1.21 | −0.89 | −1.51 | −0.41 | −0.08 |

*Derived by means of a combination of the CERES EBM result with the ERA5 temperature anomaly; see materials and methods.

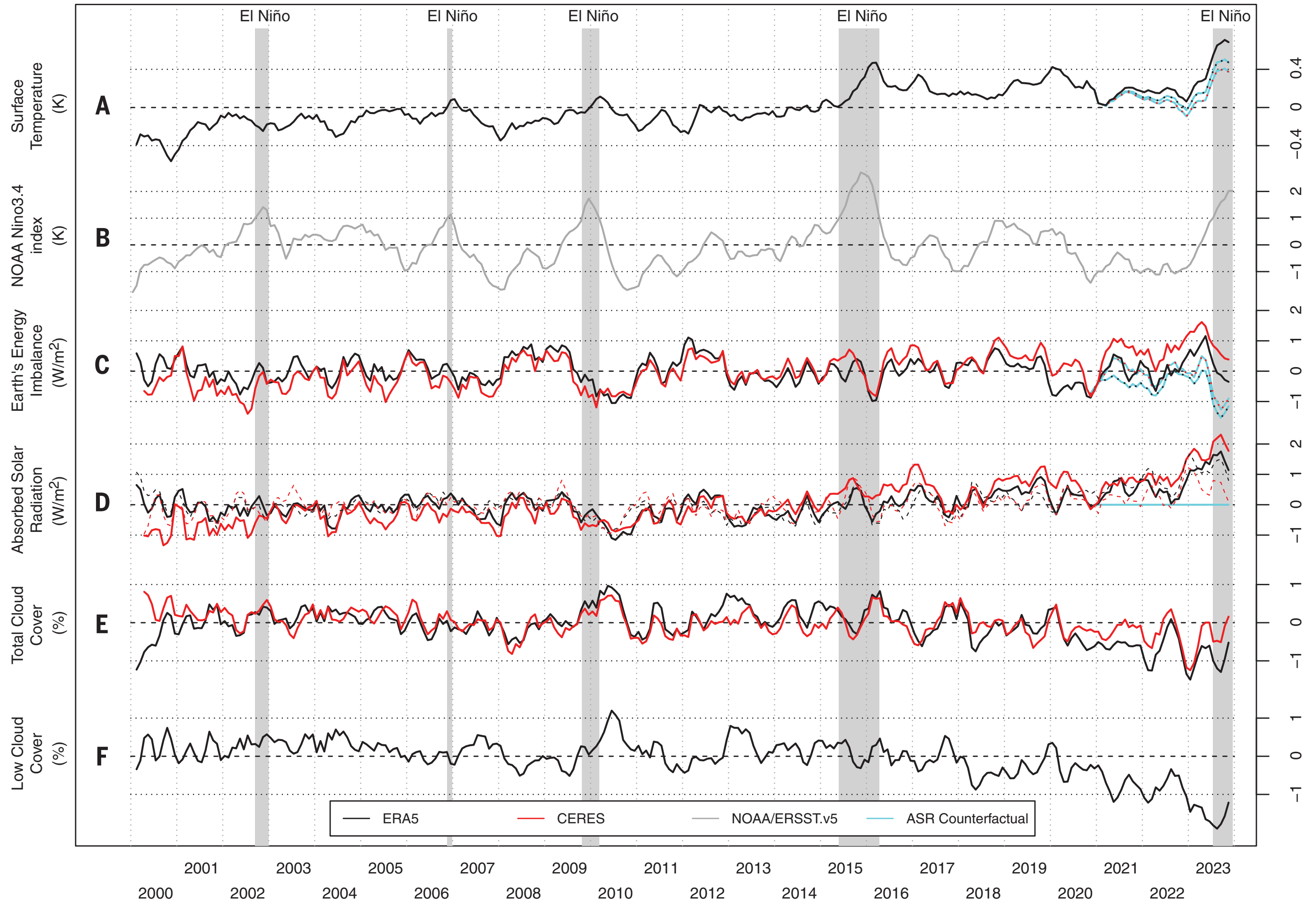


**Fig. 2. Global mean anomalies of key parameters related to Earth's temperature, energy budget, and clouds.** Three-month running-mean anomalies relative to 2001–2022 of (**A**) surface (skin) temperature, (**B**) National Oceanic and Atmospheric Administration (NOAA) Ocean Niño 3.4 index, (**C**) Earth's TOA total energy imbalance, (**D**) TOA net solar radiation (or ASR), (**E**) TCC fraction, and (**F**) low-cloud cover fraction. Red curves show satellite data from CERES, and black curves show reanalysis data from ERA5. Dashed curves in (D) show the TOA solar cloud radiative effect inferred from TCC anomalies (CREtc). Cyan curves show the full counterfactuals based on a two-layer EBM in which ASR anomalies are assumed to be zero from the beginning of December 2020 onward. El Niño periods with anomalies exceeding +1 K are highlighted with gray shading. Annual means for ERA5 data starting at 1940 and additional quantities are shown in figs. S1 and S2.

trends (Fig. 4, C and D) and negative cloud anomalies and trends (fig. S5, A and B) are highly correlated. However, the actual influence of TCC on the shortwave cloud radiative effect (CRE) and thus ASR depends strongly on surface albedo and TOA incident solar radiation. To quantify the contribution of changes in TCC and covarying optical depth to ASR anomalies, we derived empirical relations between TCC and ASR anomalies for CERES and ERA5 based on 2001–2014 data, when ASR was still relatively stationary (materials and methods).

The shortwave CRE inferred from total cloud cover (CREtc) closely resembles spatial ASR patterns (fig. S5, C and D, versus fig. S4, C and D), with the exception of the polar regions, where surface albedo anomalies linked to sea ice and snow can dominate ASR anomalies. In ERA5 (fig. S6, C and D versus G and H), not just the patterns but also the magnitudes match closely, with the 2023 global mean CREtc of +1.21 W m$^{-2}$ being close to the ASR anomaly of +1.33 W m$^{-2}$ (Table 1) and with coherent temporal variations (Fig. 2D). In CERES, however, the CREtc anomalies and trends account for only about one-third of the global mean ASR anomaly and trend (Table 1). This is consistent with the combination of a stronger ASR trend and weaker TCC trend in CERES compared with those in ERA5, particularly since 2020 (Fig. 2, D and E).

The different degree to which CREtc anomalies contribute to ASR anomalies cannot be explained by different surface albedo contributions, which are similar between the datasets, largely constrained to high latitudes, and, overall, too small. Rather, it suggests that, according to CERES, either cloud reflectivity has increased beyond pre-2015 TCC-CRE relations, clear-sky ASR has increased beyond the influence of surface albedo, or both. Before addressing the possible role of aerosols, which can influence both cloud and clear-sky reflectivity as well as cloud amount (*11*, *36*–*38*), we considered the height-dependence of cloud trends.

### Cloud anomalies are mainly due to reduced low-level clouds

According to ERA5, the TCC trends and anomalies are related mainly to declining low-level cloud cover (Figs. 2F, 3C, and 4, E and F, and fig. S3G), whereas high- and mid-level clouds have declined only slightly, if at all (fig. S2, D and E), consistent with previous findings (*19*). Regions with coherent low-cloud reductions both over 2013–2022 and in 2023 include the warm pool region around the maritime continent and the northern extratropical western Pacific, as well as large parts of the Atlantic and

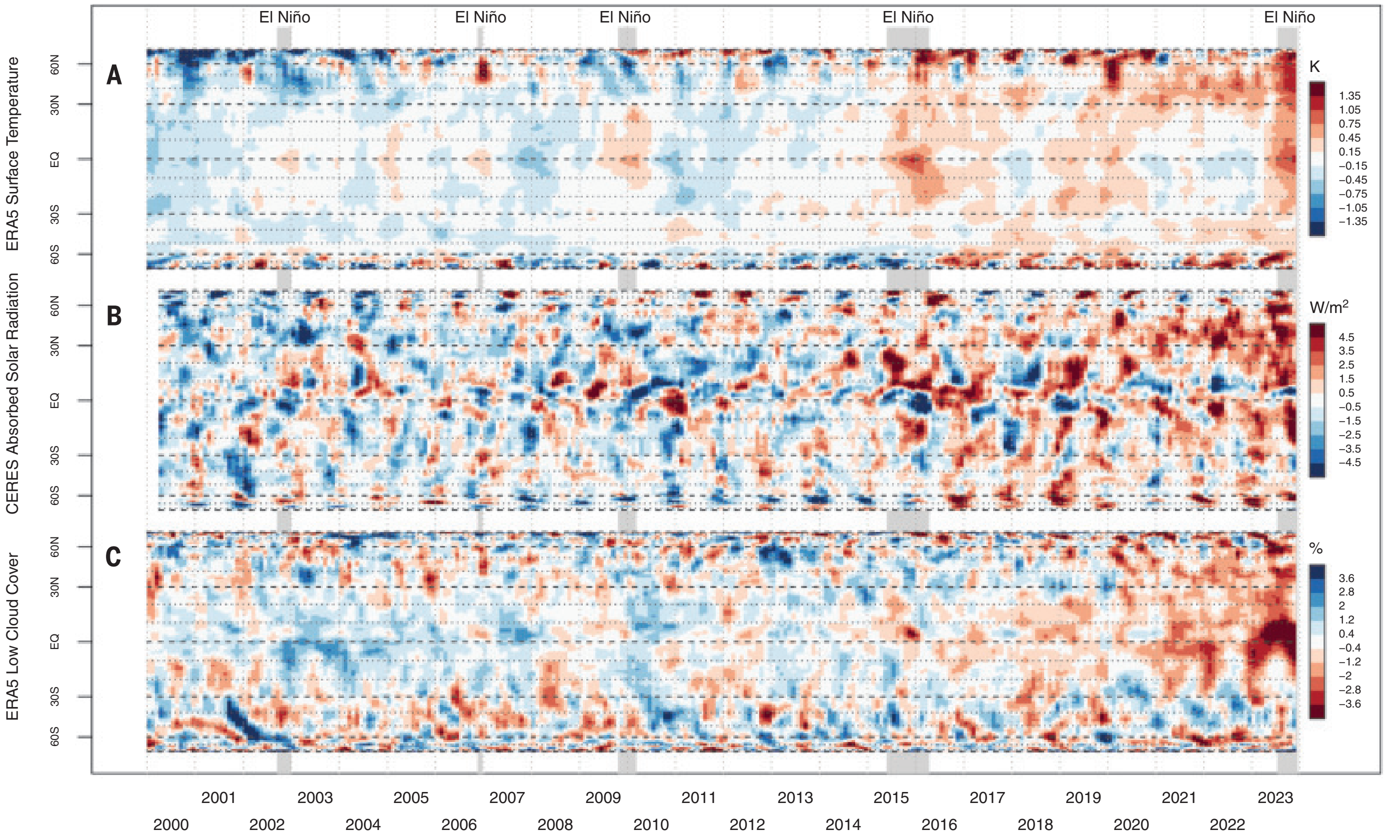


**Fig. 3. Zonal mean anomalies of key parameters related to Earth's temperature, energy budget, and clouds.** Three-monthly running-mean anomalies relative to 2001–2022 of (**A**) ERA5 surface (skin) temperature, (**B**) CERES ASR, and (**C**) ERA5 low-cloud cover. El Niño periods with anomalies exceeding +1 K are highlighted with gray shading. Latitude spacing corresponds to cos(latitude) for an equal-area representation. Additional parameters are shown in fig. S4.

adjacent land regions. Most of these regions also exhibited reduced TCC (figs. S5, A and B, and S6, A and B) and thus increased ASR (Fig. 4, C and D, and fig. S6, G and H), with the exception of the eastern warm pool, where changes in higher-level clouds have overcompensated for the low-cloud reductions.

The South Atlantic between 20°S and the equator did not exhibit negative low-cloud anomalies in 2023 but rather a pronounced trend prior to 2023. Also, in large parts of the East Pacific and South Indian Ocean, the 2023 low-cloud anomalies were not consistent with the 2013–2022 trend. This suggests that interannual variability has dominated the anomalies and trends there, even though corresponding patterns in the composite of El Niño onset years (fig. S6E) are weaker.

Globally averaged, the negative low-cloud anomaly in 2023 was about −1.5%, following a decadal trend of −1.27% per decade (Table 1). Relative to the absolute global mean low-cloud cover around 38% (fig. S1J), the 2023 anomaly amounted to a surprising −4.0%. It is notable that not only did the eastern North Atlantic, one of the main drivers of the GMST surge [Fig. 4B; (*5*)], experience pronounced low-cloud reductions in 2023 (Fig. 4F), but almost the entire Atlantic experienced a substantial decline over the previous decade (Fig. 4E).

Further characterizing the cloud anomalies in and beyond 2023 based on more detailed cloud data, such as Moderate Resolution Imaging Spectroradiometer (MODIS)–derived cloud properties (*15*, *19*, *39*), will be important to assess the cloud trends and to further resolve differences between CERES and ERA5. This includes the inconsistent total cloud trends, which may hint at an overestimation of the low-cloud cover decline in ERA5 that is related to the reliance on model physics and unrealistic aerosol handling. More detailed analyses will also be required to understand the causes of the observed cloud and ASR anomalies and trends, including the role of aerosols.

## The role of aerosols remains unclear

There is evidence for sizable aerosol contributions to recent warming trends (*11*), but isolating the contribution of indirect aerosol effects to cloud amount and reflectivity changes remains challenging. By contrast, the clear-sky ASR can provide evidence for direct aerosol effects, despite confounding influences from surface albedo and atmospheric water vapor (*40*). The sea ice and snow retreat has led to increased global mean clear-sky ASR (figs. S2A and S3H) and dominates clear-sky ASR anomalies at high latitudes (figs. S5, E and F, and S6, E and F). Over the open ocean, however, clear-sky ASR anomalies and trends in CERES (fig. S5, E and F) may hint at direct aerosol effects, albeit these are an order of magnitude smaller than the corresponding all-sky values (Fig. 4, C and D).

In 2023, CERES clear-sky ASR anomalies were broadly positive between the equator and about 45°N over the Atlantic and Pacific Oceans, with peaks off the East Asian Pacific and African Atlantic coasts (fig. S5F). The latter may be related to reduced transport of Saharan dust owing to weakened trade winds during the northern spring and summer of 2023 (*5*), whereas the former and the broader positive signal may hint at reduced aerosols of different origin, possibly as a result of reduced sulfur emissions from shipping (*9*, *41*). Given that changed aerosol concentrations are a prerequisite for indirect aerosol effects, this also suggests that aerosols may have contributed to the reduced cloud cover and/or cloud reflectivity in these regions (*9*, *41*). However, whereas the 2013–2022 clear-sky ASR trends suggest a consistent increase along the East Asian Pacific coast, trends

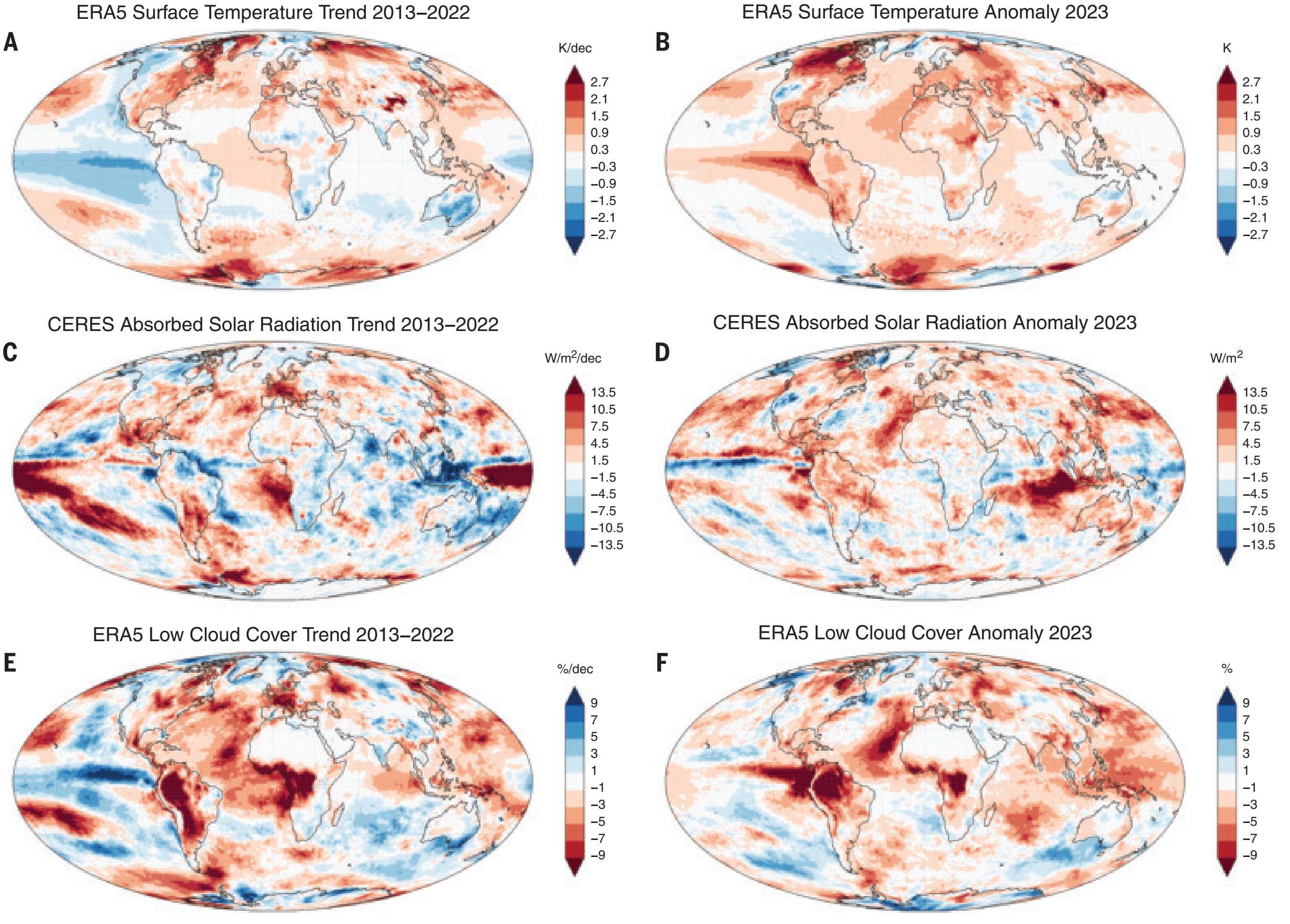


**Fig. 4. Decadal 2013–2022 trends and annual mean 2023 anomalies of key parameters related to Earth's temperature, energy budget, and clouds.** Trends and anomalies relative to 2001–2022 of (**A** and **B**) ERA5 surface (skin) temperature, (**C** and **D**) CERES ASR and (**E** and **F**) ERA5 low-cloud cover fraction. Additional parameters are shown in figs. S5 and S6.

prior to 2023 were rather negative over the Atlantic and most of the Southern Hemisphere oceans (fig. S5E). The weak negative ISR trend over this period (Table 1) is insufficient to explain these negative clear-sky ASR trends. The contribution of potentially reduced aerosol effects associated with the shipping regulations in 2015 and 2020 thus remains unclear.

Apart from the regions with sea ice and snow retreat, clear-sky ASR anomalies are much weaker and smoother in ERA5 (fig. S6, E and F), in which aerosols are prescribed (*17*) based on forcing data of the Coupled Model Intercomparison Project phase 5 (CMIP5). Alongside the recent ISR underestimation, this may add to the lower 2023 ASR and imbalance anomalies in ERA5 relative to CERES (Table 1), particularly over the North Pacific and North Atlantic.

### Temperature response to reduced planetary albedo can explain the warming gap

Earth's climate responds to different types of forcing in a complex way (*42*), but it is possible to estimate the influence of the recent ASR anomalies on GMST with a two-layer energy balance model [EBM; (*43*, *44*)]. We constructed counterfactuals in which ASR anomalies are assumed to be zero from December 2020 onward, which was about the last time that ASR anomalies were close to zero and when the clearest low-cloud and ASR trends set in (Fig. 2, D and F). We integrated monthly CERES and ERA5 ASR anomalies relative to 2001–2022 until December 2023 (materials and methods and figs. S10 and S11). Subtracting the EBM upper-layer temperature response from the observed GMST provides counterfactual realizations of how GMST may have evolved without the ASR anomalies; counterfactuals of the imbalance (EEI) are constructed in the same way. Additional counterfactuals are constructed with ASR anomalies only from different zonal bands and from incident solar radiation. Considering only the shortwave perturbations is reasonable, given that the planetary albedo decline is associated primarily with low-level clouds, which lack the compensating longwave effects of mid- and high-level clouds (*45*).

Based on the full counterfactuals, the 2023 annual mean GMST may have been 0.25 ± 0.05 K cooler based on the CERES counterfactual and 0.16 ± 0.03 K cooler based on the ERA5 counterfactual (cyan curves in Fig. 2A). Subtracting 0.027 ± 0.005 K because of the solar intensity increase after December 2020 (fig. S2F) in CERES, which is not captured in ERA5 (−0.016 ± 0.003 K), reduces the discrepancy. The full counterfactual total imbalance (EEI) anomaly drops below −1 W m$^{-2}$ toward the end of 2023 (Fig. 2C). Adjusting again for the solar intensity results in less-negative counterfactual EEI anomalies in both cases, which is more consistent with previous El Niño events. Importantly, the CERES estimate for the effect of planetary albedo alone on the 2023 GMST is +0.22 ± 0.04 K (Fig. 1B), which is approximately the magnitude of the unexplained warming.

The polar regions beyond 55°S and 55°N, where surface albedo decline due to sea ice and snow retreat strongly affects the ASR, jointly contribute +0.030 ± 0.006 K (12%) to the ASR-driven warming response (Fig. 1B). The remaining 0.19 ± 0.04 K are dominated by the northern mid-latitudes and tropics, where cloud changes drive ASR anomalies. By comparison, based on the Niño 3.4 index in 2023 compared with that of previous El Niño onset years and their annual mean GMST anomalies (materials and methods), El Niño has contributed only +0.07 ± 0.04 K to the 2023 temperature anomaly. This leaves a residual of +1.16 ± 0.09 K warming that may have occurred in 2023 without the anomalous ASR and El Niño (Fig. 1B and fig. S8), including the bulk of anthropogenic warming and any additional factors such as longwave effects from the Hunga Tonga–Hunga Ha'apai eruption. The total GMST response of +0.22 ± 0.04 K to the reduced planetary albedo is broadly consistent with the amount of cooling observed after major volcanic eruptions relative to the respective forcing [fig. S3, A and D; (*42*, *46*)].

## Potential implications for climate sensitivity

Three fundamental mechanisms may have contributed to the record-low planetary albedo associated with reduced low-level clouds: internal variability, an emerging low-cloud feedback, and aerosol effects. Contributions from internal variability would subside and leave our expectation of the longer-term warming unaffected. The relative stationarity of the low-cloud cover until about 2015 (Figs. 2F and 3C) speaks against short-term variability, but longer-term variability associated with, for example, the Atlantic multidecadal variability (*5*, *47*) could contribute to the observed trends, also given that ocean surface warming can reduce low-cloud cover (*35*, *48*, *49*).

The latter mechanism is also essential if the recent trends are due to an emerging low-cloud feedback that is unrelated to internal variability, complicating a separation of the two. The response of low clouds is the largest source of uncertainty driving differences in climate sensitivity between climate models (*50*). This holds even after the expected range of low-cloud response and climate sensitivity could be reduced with observational constraints (*35*, *51*), giving an assessed range of combined marine low-cloud feedback of $+0.37 \pm 0.33$ W m$^{-2}$ K$^{-1}$ [(*51*); materials and methods]. If a substantial low-cloud feedback that is closer to the upper end of this range now emerges in observations, the lower end of realistic climate sensitivity estimates of 2.3 to 4.7 K (*51*) may need to be adjusted upward.

The average combined shortwave forcing by aerosols in CMIP6 climate models was −1.26 W m$^{-2}$ (*52*). A near-complete loss of anthropogenic aerosols would thus be required to match the observed ASR anomaly in 2023, suggesting either an underestimated aerosol effect in models (*9*) or strong contributions from internal variability or low-cloud feedback (*19*). Even though the negative correlation between the aerosol effect and climate sensitivity found in CMIP3 (*53*) was weaker in later CMIPs (*52*, *54*), a stronger historical aerosol cooling would require a higher sensitivity to greenhouse forcing to reproduce the observed temperature record.

If the cloud-related albedo decline was not largely caused by internal variability, the 2023 extra heat may be here to stay, and Earth's climate sensitivity may be closer to the upper range of present estimates. We may thus be closer to the temperature targets defined in the Paris Agreement than previously thought, with potentially strong implications for remaining carbon budgets.

## REFERENCES AND NOTES

1. T. Iwakiri *et al*., *Geophys. Res. Lett.* **50**, e2023GL105763 (2023).
2. Copernicus Climate Change Service (C3S), European State of the Climate 2023, Summary (2024); https://doi.org/10.24381/bs9v-8c66.
3. G. Schmidt, *Nature* **627**, 467 (2024).
4. T. Rackow *et al*., *Nat. Commun.* **13**, 637 (2022).
5. T. Kuhlbrodt, R. Swaminathan, P. Ceppi, T. Wilder, *Bull. Am. Meteorol. Soc.* **105**, E474–E485 (2024).
6. L. A. Roach, W. N. Meier, *Nat. Rev. Earth Environ.* **5**, 235–237 (2024).
7. G. Kopp, *J. Space Weather Space Clim.* **6**, A30 (2016).
8. M. R. Schoeberl *et al*., *J. Geophys. Res. Atmos.* **129**, e2024JD041296 (2024).
9. J. Hansen *et al*., *Oxford Open Climate Change* **3**, kgad008 (2024).
10. P. Manshausen, D. Watson-Parris, M. W. Christensen, J. P. Jalkanen, P. Stier, *Nature* **610**, 101–106 (2022).
11. Ø. Hodnebrog *et al*., *Commun. Earth Environ.* **5**, 166 (2024).
12. H. Wang *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2313797121 (2024).
13. T. Yuan *et al*., *Commun. Earth Environ.* **5**, 281 (2024).
14. N. G. Loeb *et al*., *J. Clim.* **31**, 895–918 (2018).
15. P. Minnis et al, in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59 (IEEE, 2021), pp. 2744–2780.
16. T. Dudok de Wit, G. Kopp, C. Fröhlich, M. Schöll, *Geophys. Res. Lett.* **44**, 1196–1203 (2017).
17. H. Hersbach *et al*., *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
18. B. Bell *et al*., *Q. J. R. Meteorol. Soc.* **147**, 4186–4227 (2021).
19. N. G. Loeb *et al*., *Surv. Geophys.* (2024).
20. H. Song, Y.-S. Choi, H. Kang, Research Square 10.21203/rs.3.rs-3781529/v1 [Preprint] (2024); https://www.researchsquare.com/article/rs-3781529/v1.
21. N. G. Loeb *et al*., *Geophys. Res. Lett.* **48**, e2021GL093047 (2021).
22. A. Donohoe, K. C. Armour, A. G. Pendergrass, D. S. Battisti, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 16700–16705 (2014).
23. Y. He, K. Wang, C. Zhou, M. Wild, *Geophys. Res. Lett.* **45**, 4281–4289 (2018).
24. S. Kato, *J. Clim.* **22**, 4893–4907 (2009).
25. C. Radley, S. Fueglistaler, L. Donner, *J. Clim.* **27**, 3100–3113 (2014).
26. G. L. Stephens *et al*., *Rev. Geophys.* **53**, 141–163 (2015).
27. A. Timmermann *et al*., *Nature* **559**, 535–545 (2018).
28. B. Huang *et al*., *J. Clim.* **30**, 8179–8205 (2017).
29. A. Letterly, J. Key, Y. Liu, *Cryosphere* **12**, 3373–3382 (2018).
30. A. Riihelä, R. M. Bright, K. Anttila, *Nat. Geosci.* **14**, 832–836 (2021).
31. A. Donohoe, D. S. Battisti, *J. Clim.* **24**, 4402–4418 (2011).
32. N. G. Loeb *et al*., *J. Clim.* **32**, 5003–5019 (2019).
33. S. Kato *et al*., *Geophys. Res. Lett.* **33**, L19804 (2006).
34. M. D. Zelinka, S. A. Klein, D. L. Hartmann, *J. Clim.* **25**, 3736–3754 (2012).
35. S. A. Klein, A. Hall, J. R. Norris, R. Pincus, *Surv. Geophys.* **38**, 1307–1329 (2017).
36. S. Twomey, *J. Atmos. Sci.* **34**, 1149–1152 (1977).
37. J. Quaas, B. Stevens, P. Stier, U. Lohmann, *Atmos. Chem. Phys.* **10**, 6129–6135 (2010).
38. J. Li *et al*., *Nat. Rev. Earth Environ.* **3**, 363–379 (2022).
39. G. Tselioudis, W. B. Rossow, F. Bender, L. Oreopoulos, J. Remillard, *Clim. Dyn.* **62**, 9319–9332 (2024).
40. K. E. Taylor *et al*., *J. Clim.* **20**, 2530–2543 (2007).
41. M. S. Diamond, *Atmos. Chem. Phys.* **23**, 8259–8269 (2023).
42. J. Hansen *et al*., *J. Geophys. Res.* **110**, D18104 (2005).
43. I. M. Held *et al*., *J. Clim.* **23**, 2418–2427 (2010).
44. O. Geoffroy *et al*., *J. Clim.* **26**, 1841–1857 (2013).
45. M. D. Zelinka, C. Zhou, S. A. Klein, *Geophys. Res. Lett.* **43**, 9259–9269 (2016).
46. D. Zanchettin *et al*., *Geosci. Model Dev.* **15**, 2265–2292 (2022).
47. K. Li, F. Zheng, J. Zhu, Q.-C. Zeng, *Adv. Atmos. Sci.* **41**, 1017–1022 (2024).
48. C. L. Boehm, D. W. J. Thompson, *J. Clim.* **36**, 2753–2762 (2023).
49. M. Athanase, A. Sánchez-Benítez, H. F. Goessling, F. Pithan, T. Jung, *Commun. Earth Environ.* **5**, 53 (2024).
50. X. Qu, A. Hall, S. A. Klein, P. M. Caldwell, *Clim. Dyn.* **42**, 2603–2626 (2014).
51. S. C. Sherwood *et al*., *Rev. Geophys.* **58**, RG000678 (2020).
52. C. J. Smith *et al*., *Atmos. Chem. Phys.* **20**, 9591–9618 (2020).
53. J. T. Kiehl, *Geophys. Res. Lett.* **34**, L22710 (2007).
54. P. M. Forster *et al*., *J. Geophys. Res. Atmos.* **118**, 1139–1150 (2013).

## ACKNOWLEDGMENTS

We thank H. Hersbach and A. Simmons for answering our questions about the interpretation of ERA5 trends and aerosol handling. S. Bathiany, T. Becker, S. Cheedela, and A. Boetius provided valuable comments on a draft version of this paper. We thank the Copernicus Climate Change Service (C3S) and the National Aeronautics and Space Administration (NASA) for the development, production, and provision of the CERES-EBAF and ERA5 datasets. We thank the German Climate Computing Centre (DKRZ) for providing simple access to the ERA5 data and computing resources for the data analysis. We also thank two anonymous reviewers for thoughtful remarks that helped improve this work. **Funding:** H.F.G. and T.J. are supported by internal funds of the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, which is part of the Helmholtz Association and funded by the Germany Ministry of Education and Research (BMBF). This research was supported by the European Commission Horizon 2020 Framework Programme nextGEMS (grant no. 101003470). **Author contributions:** Conceptualization: H.F.G.; Methodology: H.F.G., T.R.; Investigation: H.F.G., T.R., T.J.; Visualization: H.F.G., T.R.; Writing – original draft: H.F.G., T.R., T.J.; Writing – review and editing: H.F.G., T.R., T.J. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** ERA5 data are available from the Copernicus Climate Change Service (C3S) Climate Data Storage at https://cds.climate.copernicus.eu/. CERES-EBAF data are available from NASA at https://asdc.larc.nasa.gov/project/CERES. NOAA Climate Prediction Center Ocean Niño Index data are available at https://www.cpc.ncep.noaa.gov/data/indices/oni.ascii.txt.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adq7280
Materials and Methods
Figs. S1 to S11
Table S1
References

Submitted 30 May 2024; accepted 19 November 2024
Published online 5 December 2024
10.1126/science.adq7280

SYNTHETIC BIOLOGY

# Engineering synthetic phosphorylation signaling networks in human cells

**Xiaoyu Yang[1,2,3], Jason W. Rocks[4], Kaiyi Jiang[1]†, Andrew J. Walters[1,3,5,6], Kshitij Rai[1,3], Jing Liu[1], Jason Nguyen[1], Scott D. Olson[6], Pankaj Mehta[4,7,8], James J. Collins[9,10,11], Nichole M. Daringer[12]*, Caleb J. Bashor[1,3,13]***

**Protein phosphorylation signaling networks have a central role in how cells sense and respond to their environment. We engineered artificial phosphorylation networks in which reversible enzymatic phosphorylation cycles were assembled from modular protein domain parts and wired together to create synthetic phosphorylation circuits in human cells. Our design scheme enabled model-guided tuning of circuit function and the ability to make diverse network connections; synthetic phosphorylation circuits can be coupled to upstream cell surface receptors to enable fast-timescale sensing of extracellular ligands, and downstream connections can regulate gene expression. We engineered cell-based cytokine controllers that dynamically sense and suppress activated T cells. Our work introduces a generalizable approach that allows the design of signaling circuits that enable user-defined sense-and-respond function for diverse biosensing and therapeutic applications.**

Cells universally use protein phosphorylation signaling networks to adapt to chemical and physical cues from their external environment. In metazoan cells, these networks consist of multilayered pathways that rapidly and reversibly convert signals detected by cell surface receptors into diverse responses such as cell movement, secretion, metabolism, and gene expression (*1*). The ability to design artificial phospho-signaling circuits that exhibit native-like signaling behavior, yet can be programmed with custom-defined input-output connectivity, could be used to create powerful biotechnology (*2*), including human cell–based therapeutics that autonomously sense and respond to specific physiological signals or disease markers on a fast timescale (*3*, *4*).

Despite this potential, engineering of phospho-signaling circuits has lagged behind that of genetic circuits (*5*), for which advances in both microorganismal and mammalian settings have been enabled by design frameworks that leverage the intrinsic modularity of promoters and coding regions (*6*), as well as carefully benchmarked sets of genetic parts (*7*, *8*)—features that have facilitated scaling of circuit complexity and fine tuning of circuit behavior with predictive quantitative models (*9*, *10*). Progress to date in engineering phospho-signaling circuits has included development of two-component phosphorylation pathways as programmable sense-and-respond modules in bacteria (*11*–*13*) and mammalian cells (*14*–*16*) and complex fast-timescale phosphorylation circuitry in yeast (*17*–*19*). In human cells, rewiring of native phospho-signaling networks has been used to create compact therapeutic programs (*20*) and sense-and-respond circuits that connect surface receptors to transcriptional outputs (*21*, *22*). We sought to facilitate the de novo design of multilayered synthetic phosphorylation circuitry in human cells with programmable input-output connectivity and signal processing.

To establish such a framework, we emulated fundamental design features of native phospho-signaling networks, which are organized as sets of interlinked phosphorylation cycles comprising kinase and phosphatase activities that mutually act on a protein substrate (*23*, *24*) (Fig. 1A). Changes in cycle equilibrium occur rapidly in response to input signals and are quickly reversed upon input removal, enabling cells to adapt to environmental changes on timescales of seconds to minutes. These features motivated us to design tunable, interconnectable phosphorylation cycles as elementary units for constructing synthetic signaling circuits. To accomplish this, we took advantage of the intrinsic structural modularity of signaling proteins, which are typically composed of discrete domains that either carry out catalytic function (for example, kinase or phosphatase domains) or specify interactions with other signaling components [for example, PDZ or Src homology 3 (SH3) domains] (*25*). The recombination of catalytic and interaction domains is thought to allow rewiring signaling network connectivity during evolution (*26*) and has proved to be a versatile tool for engineering synthetic signaling network linkages (*27*) to create pathways with altered input-output relationships (*28*) or to introduce new information processing function (*20*, *29*). We reasoned that orthogonal interaction domains could be used to mutually direct kinase and phosphatase domains to act on protein substrate targets, establishing synthetic cycles of phosphorylation and dephosphorylation that operate separately from native signaling networks.

To validate this design strategy, we developed a protein domain part set that included engineered catalytic domains derived from components of immune phosphotyrosine (pY) signaling pathways (figs. S1 and S2) (*30*). These proteins naturally use recruitment-dependent mechanisms of signaling activation and are weakly expressed in most nonimmune cell types (*30*). As an initial test, we fused various pY kinase domains (fig. S3) to leucine zippers (LZs)—small, highly specific heterodimerizing protein interaction domains with tunable interaction affinities (*31*)—to create synthetically targeted kinases (synKins) (fig. S4). Synthetic substrate proteins (synSubs) for the synKins were constructed by fusing cognate LZs with ITAMs (immunoreceptor tyrosine-based activation motifs), conserved pY motifs involved in immune signaling pathway activation (Fig. 1B and fig. S4) (*32*). Plasmid constructs encoding pairs of epitope-tagged synKins and synSubs (fig. S4) were transfected into human embryonic kidney (HEK) 293T cells, and multicolor flow cytometry was used to simultaneously measure component expression (staining of epitopes) and synSub phosphorylation (staining of phosphorylated ITAMs) in single cells (fig. S5). To optimize synKin function, we tested numerous kinase domain boundaries and point mutations (fig. S6). We identified variants that showed strong expression and demonstrated phosphorylation activity toward synSub that was highly dependent on LZ-mediated kinase domain recruitment, as evidenced by nonbinding, catalytically inactive, and unrecruited (noncognate LZ) controls all showing little or no phosphorylation (Fig. 1B, right, and fig. S7, A to C). We observed minimal cross-talk of these components with host cell machinery: Expression of synKin showed limited nonspecific Y phosphorylation, and synSub phosphorylation in the absence of synKin expression was negligible (fig. S7, B and C). Further, we found that expression of our proteins showed no major limiting effect on cell viability or growth rate (fig. S7D).

We tested whether we could tune the activity of synKin toward synSub by altering the molecular properties of the domain parts. We constructed several sets of synKin variants: LZ

---

[1]Department of Bioengineering, Rice University, Houston, TX, USA. [2]Graduate Program in Systems, Synthetic and Physical Biology, Rice University, Houston, TX, USA. [3]Rice Synthetic Biology Institute, Rice University, Houston, TX, USA. [4]Department of Physics, Boston University, Boston, MA, USA. [5]Graduate Program in Bioengineering, Rice University, Houston, TX, USA. [6]Department of Pediatric Surgery, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX, USA. [7]Biological Design Center, Boston University, Boston, MA, USA. [8]Faculty of Computing and Data Science, Boston University, Boston, MA, USA. [9]Institute for Medical Engineering and Science, Department of Biological Engineering, and Synthetic Biology Center, Massachusetts Institute of Technology, Cambridge, MA, USA. [10]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [11]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA. [12]Department of Biomedical Engineering, Rowan University, Glassboro, NJ, USA. [13]Department of Biosciences, Rice University, Houston, TX, USA.
***Corresponding author. Email: caleb.bashor@rice.edu (C.J.B.); daringer@rowan.edu (N.M.D.)**
†Present address: Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA.

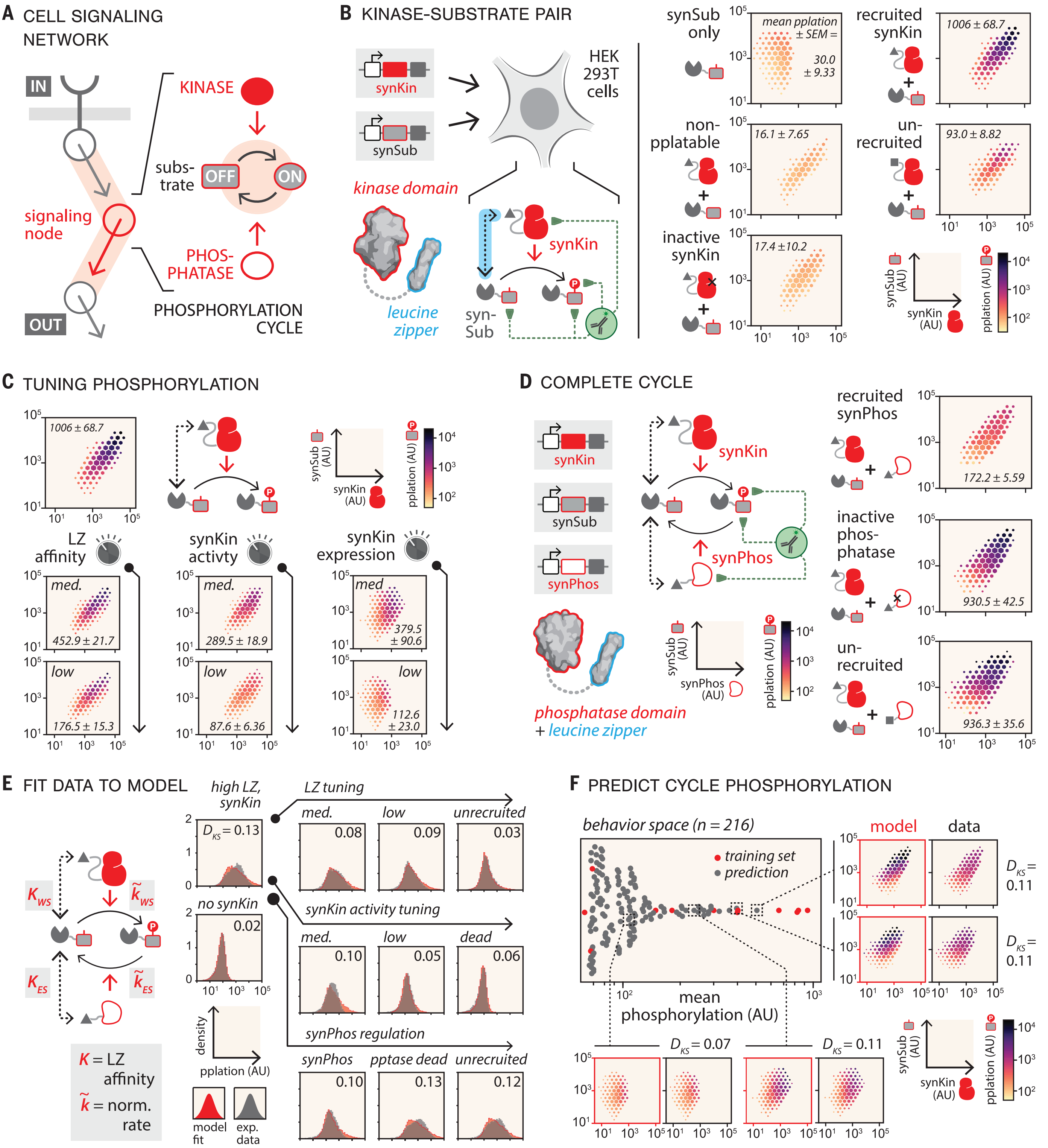


**Fig. 1. Building and tuning synthetic phosphorylation cycles in human cells.** (**A**) Phosphorylation cycles, in which kinase and phosphatase activities mutually regulate (red arrows) phosphorylation equilibrium (black arrows) of a substrate, are fundamental units that make up phospho-signaling networks. (**B**) Engineering synthetic kinase (synKin) and substrate (synSub) pairs. Plasmids (gray rectangles) encoding synKin and synSub proteins are transfected into HEK293T cells and measured for expression and phosphorylation with immunofluorescence flow cytometry (green dotted lines) after 36 hours. Leucine zippers (LZs) mediate interactions (cyan dashed line) between synKin and synSub proteins (left). Expression and phosphorylation data for synKin-synSub allele combinations are shown at right as hexagonal-hit-and-heat (HHH) plots. Through the whole study, the expression space was uniformly binned into grids; the hexagon size indicates the cell counts in each bin, the largest hexagon size indicates the highest cell density, and the smallest hexagon size indicates the lowest density, with hexagon sizes in between indicating log10-normalized cell counts proportionally (fig. S5). Values associated with HHH plots are mean phosphorylation (AU, arbitrary units) ± SEM ($n$ = 3 biological replicates). (**C**) Tuning synSub phosphorylation. Phosphorylation was measured for synKin-synSub

compositions that feature parts that tune LZ binding affinity, synKin expression level, and synKin activity. (Top) The behavior of the default synKin-synSub from (B) for comparison. Numbers are associated with HHH plots, mean phosphorylation (AU) ± SEM ($n = 3$ biological replicates). (**D**) Complete synthetic phosphorylation cycle. (Left) LZ-recruited synPhos dephosphorylates synSub. (Right) Values adjacent to HHH plots of synPhos variants indicate mean phosphorylation (AU) ± SEM ($n = 3$ biological replicates). (**E**) Modeling synthetic phosphorylation cycle equilibrium. (Left) LZ interactions [$K_{WS}$, synKin-synSub affinity (W, writer; S, substrate); $K_{ES}$, synPhos-synSub affinity (E, eraser; S, substrate)] and component activity ($\tilde{k}_{WS}$, normalized phosphorylation rate; $\tilde{k}_{ES}$, normalized dephosphorylation rate) of the cycle (model description is available in the supplementary text). Phosphorylation cycle data from (B) to (D) were used for fitting. Model-predicted phosphorylation distributions (red) are plotted against experimentally measured distributions (gray). Kolmogorov-Smirnov divergence ($D_{KS}$) values comparing model with experiment are shown for each plot. norm., normalized. (**F**) Model-predicted phosphorylation cycle behavior space. The beeswarm plot shows phosphorylation values for predicted (gray dots) and training (red dots) part compositions (216 total). HHH plots for compositions from across design space (indicated with black dotted lines) compare predicted phosphorylation (red borders) and measured phosphorylation (black borders). Compositions of highlighted circuits are identical to synPhos composition in (D) (strong synKin and synPhos LZ, strong synKin and synPhos activity, and strong synKin and synPhos expression) except for the following: bottom left, low synKin expression, dead synPhos; bottom middle, medium synKin expression, weak synPhos LZ; top right, low synPhos; bottom right, weak synPhos activity, medium synKin LZ.

sequence variants were introduced to tune binding affinity to synSub (*31*); catalytic turnover rate was adjusted with previously reported pY kinase domain active site mutations (fig. S4) (*33*); and expression level was tuned by introducing Kozak sequence variants to synKin expression constructs, resulting in differential rates of protein translation (*34*). When we tested each synKin part set, we observed modulation of synSub phosphorylation across a 10- to 20-fold range (Fig. 1C). To create complete cycles of reversible synSub phosphorylation, we developed synthetically targeted pY phosphatases (synPhos). We took an approach similar to what we used for the synKins, identifying domain variants derived from pY phosphatases involved in immune signaling (figs. S8 to S9) (*35*) and fusing them to the same LZ species as that with the synKin. When coexpressed with a synKin-synSub pair, synPhos dephosphorylated synSub in a recruitment-dependent and phosphatase activity–dependent manner (Fig. 1D). Taken together, these results validate our design strategy for constructing synthetic phosphorylation cycles and demonstrate that a simple part set that consists of catalytic and interactions domain variants can be used to rationally control intracellular phosphorylation equilibrium.

We developed a model to quantitatively describe the relationship between phosphorylation cycle equilibrium and part biophysical properties. To achieve this, we first converted single-cell fluorescence values into stoichiometric equivalents for all phosphorylation cycles depicted in Fig. 1, B to D, by normalizing different color fluorophores to an enhanced green fluorescent protein (EGFP) reference (figs. S10 and S11). These transformed data were then fit to a nonequilibrium thermodynamic model (Fig. 1E and fig. S12) to obtain part-specific parameters for LZ variant interaction affinities ($K_{WS}, K_{ES}$) and catalytic turnover rates of synKin and synPhos variants $(\tilde{k}_{WS}, \tilde{k}_{ES})$ (fig. S12). The parameterized model was then used to predict phosphorylation states for all part combinations within our design space ($n = 216$ total compositions) (Fig. 1F and fig. S13). To validate these predictions, we constructed and measured compositions from across the predicted behavior distribution; all showed excellent overall agreement with the model (Fig. 1F), demonstrating that the functional modularity inherent in our design scheme lends itself to prediction of regulated phosphorylation behavior based on individual parts properties. Last, we used the model to analyze our part set for compositions that could undergo sharp, ultrasensitive transition in phosphorylation upon changes in component activity, expression levels, and LZ affinity (fig. S14). From this analysis, we determined that the behavior for the majority of compositions in our part design space fall in a linear regime, although tuning to higher sensitivity is possible by using higher synKin LZ affinity, lower synPhos LZ affinity, and higher synSub expression.

Native signaling networks convert protein phosphorylation into molecular outputs through various mechanisms (*36*), including allosteric regulation of protein activity, changes in protein localization and stability, and formation of new protein-protein interactions. In the latter case, phospho-specific binding domains recognize phosphorylated substrate motifs, forming interactions that facilitate downstream signaling. To form circuit connections between phosphorylation cycles, we used engineered SH2 domains, which bind to pY-containing motifs and are conserved among metazoans (Fig. 2A) (*37*). We used a transcriptional reporter (fig. S15, A and B) to not only validate synKin activity–dependent SH2-pY motif interactions but also to identify part sets with orthogonal interaction specificities (fig. S15D): Tandem SH2 (tSH2) domains and an engineered multivalent SH2 respectively bound synKin-phosphorylated ITAMs (fig. S15C) (*38*) and a dual pY motif derived from the immune signaling protein SLP76 (SH2-domain-containing leukocyte protein of 76 kDa) (fig. S16A), with no observable cross-talk (fig. S16B).

To test whether these phospho-dependent interactions could be used to link two phosphorylation cycles together, we engineered a "phospho-couple" (PC) protein that integrates the functions of an upstream synSub and downstream synKin by fusing a kinase domain to three substrate motifs and placing a rigid linker domain between them (fig. S17) to limit cis-phosphorylation (Fig. 2A, right, and fig. S17). This design facilitated sequential activation of phosphorylation cycles: When we expressed a four-protein system (upstream synKin, PC, synPhos, and downstream SH2-synSub) in HEK293T cells, phosphorylation of PC by the upstream synKin led to recruitment and phosphorylation of a downstream, tSH2-fused synSub (Fig. 2B). Sequential phosphorylation was dependent on upstream synKin recruitment, PC activity, and SH2-mediated recruitment, as well as 3× substrate motif valency (Fig. 2B, right).

One important systems-level property of native phosphorylation cascades is their ability to stoichiometrically amplify weak input signals into macroscopic cellular outputs (*39*). To determine whether our two-step circuit architecture could be tuned to maximize amplification of an upstream input, we expanded our quantitative model to fit data from Fig. 2B, obtaining part-specific parameters (fig. S18) that allowed behavior predictions across two-step circuit combinatorial design space ($n = 3456$ compositions) (Fig. 2C, left). We identified a region of behavior space with compositions predicted to show a >10× fold-change in downstream synSub phosphorylation upon addition of an upstream synKin ($n = 261$ compositions). Circuits from this high-gain region were enriched for features that are consistent with stoichiometric amplification, including a low PC:SH2-synSub ratio and strong synPhos activity (fig. S19). To validate model predictions, we selected several amplifier circuit compositions from this region to experimentally measure, demonstrating the general agreement of their behavior with model predictions (Fig. 2C, right). These results indicate that our part set and predictive modeling framework can be extended to guide the design of multiphosphorylation cycle networks with programmed signal-processing properties.

Having developed approaches for building, interconnecting, and predictively tuning synthetic phosphorylation cycles, we sought to engineer surface receptors that could couple extracellular ligand binding to changes in

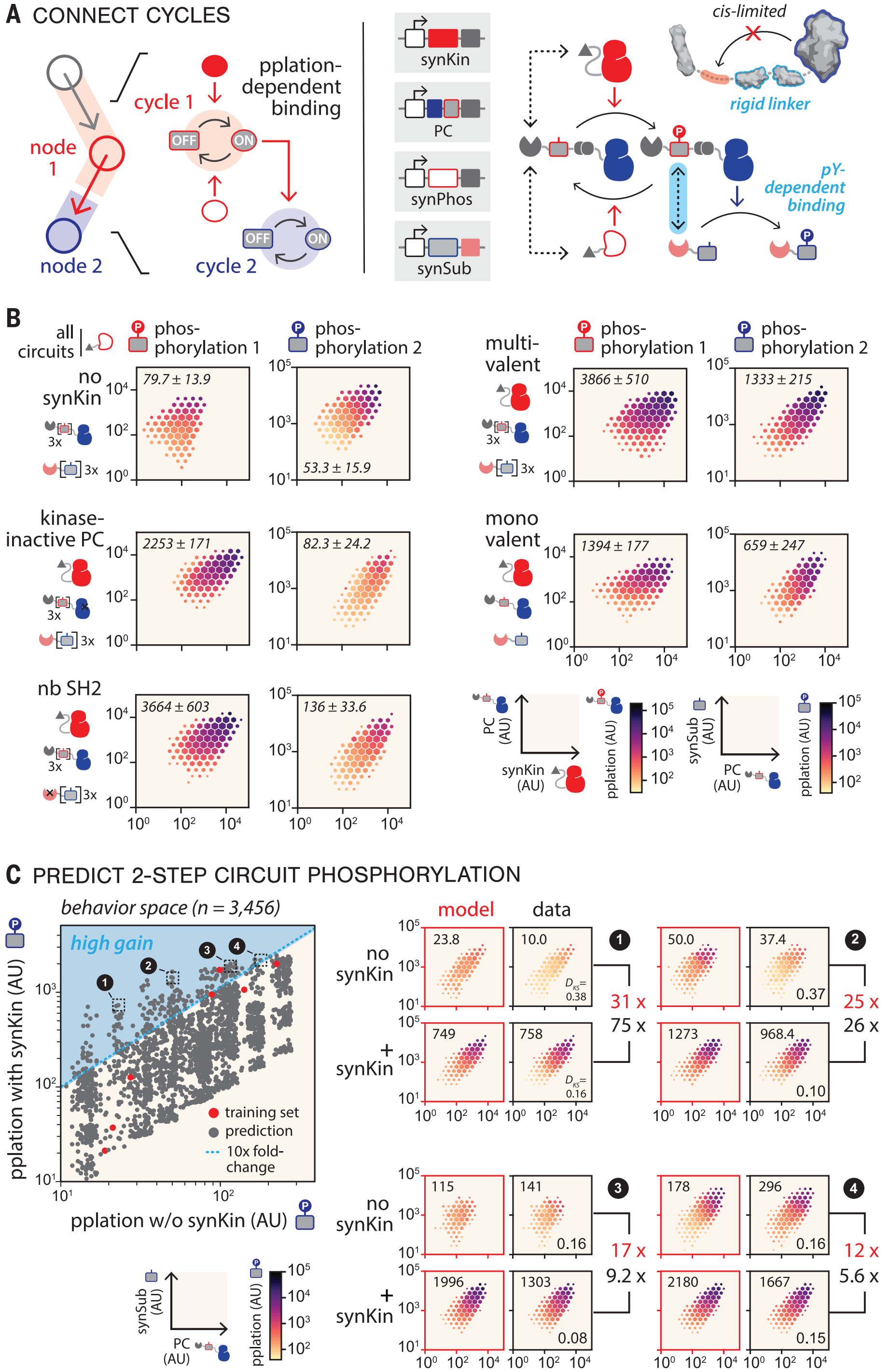


**Fig. 2. Building and tuning phosphorylation cycle networks.** (**A**) Two-step phosphorylation cascade circuit. To create a synthetic phosphorylation network connection, an upstream cycle (red) can be coupled to a downstream cycle (blue) by using a phosphorylation-dependent interaction (left). The upstream synKin phosphorylates a "phospho-couple" (PC) protein, which functions as synSub for the upstream cycle and synKin for the downstream one (right). The PC contains a rigid linker domain (cyan outline) that prevents cis substrate motif phosphorylation while supporting SH2-mediated recruitment of a downstream substrate (cyan dashed line at top right). (**B**) Two-step circuit validation. HHH plots for PC (left column) and synSub (right column) phosphorylation are shown for various circuit compositions. All circuit compositions contain synPhos. Brackets indicate the number of ITAM motif repeats. Values in each plot are mean phosphorylation (AU) ± SEM (*n* = 3 biological replicates). (**C**) Model-predicted circuit behavior space for two-step circuit. The modeling framework was used to fit steady-state phosphorylation of multivalent two-step circuits from (B) (fig. S18). Scatter plot shows predicted mean phosphorylation for compositions with (*y* axis) and without (*x* axis) upstream synKin. Red dots indicate training set compositions from fig. S18B; gray dots indicate model-predicted compositions (3456 total). Region of >10× fold-change (± synKin) is shown in cyan. Four high-gain circuits (indicated with black dotted lines) were constructed and tested for synSub phosphorylation. (Right) HHH plots for model-predicted (red border) and experimental measurements (black border) are shown. Compositions of highlighted circuits are identical with the multivalent composition in (B), except for the following: (1), medium PC kinase activity, high synPhos activity; (2), low PC expression; (3), high synKin expression, low PC expression; and (4), high PC expression, high synPhos expression. Values at the top of each plot indicate mean phosphorylation (AU) ± SEM (*n* = 3 biological replicates). $D_{KS}$ values are shown at the bottom of each experimental plot. Predicted (red) and experimentally measured (black) fold change values are shown at right of the plots.

phosphorylation cycle equilibrium (Fig. 3A, left). We constructed a pair of synthetic receptor scaffolds (figs. S20) that consisted of flexible intracellular and extracellular linker sequences and transmembrane helices. Kinase and LZ domains were appended to the cytoplasmic termini of the scaffolds, and FKBP (FK506 binding protein) and FRB* (FKBP-rapamycin-binding domain of mTOR kinase, K2095P)—domains that heterodimerize upon binding to the rapamycin analog AP21967 (*40*)—to their extracellular termini. This architecture should enable ligand-induced synSub phosphorylation as a result of receptor dimerization–enforced proximity between the synKin and LZ (Fig. 3A, right). After transfecting a four-protein system consisting of the receptor pair, synSub, and synPhos proteins, we observed a ~20-fold change in phosphorylation upon ligand addition. Circuit induction was dependent on both LZ-mediated synSub recruitment and synKin activity. Additionally, we observed that elimination of synPhos activity resulted in a lower fold-change response (6.9×), demonstrating the importance of phosphorylation cycle reversibility for optimizing circuit performance (Fig. 3B). Last, the dose response profile for the circuit was highly linear [Hill coefficient ($n_H$) = 1.1].

To assess the timescale of activation for our receptor-mediated phospho-sensor circuit, we

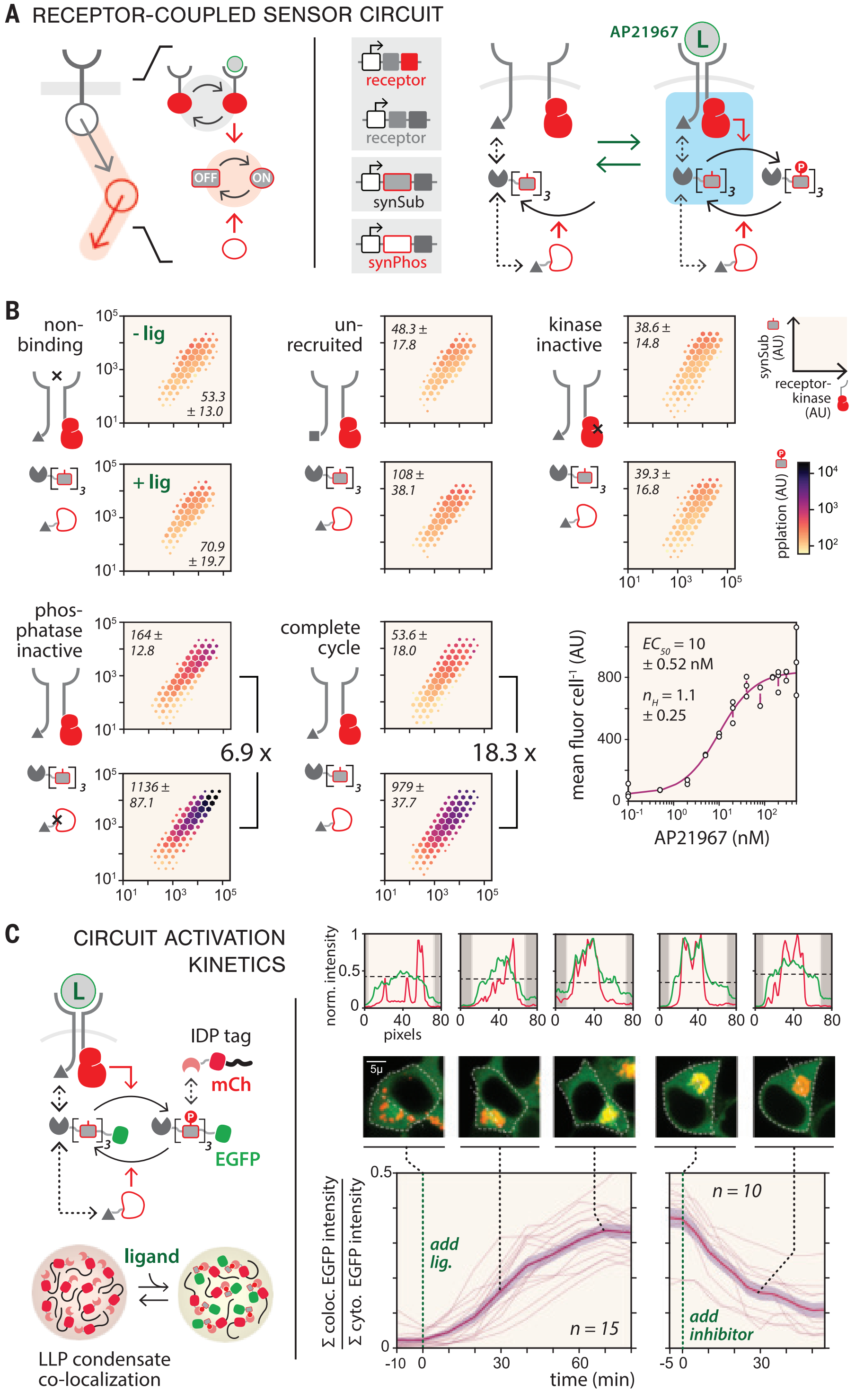
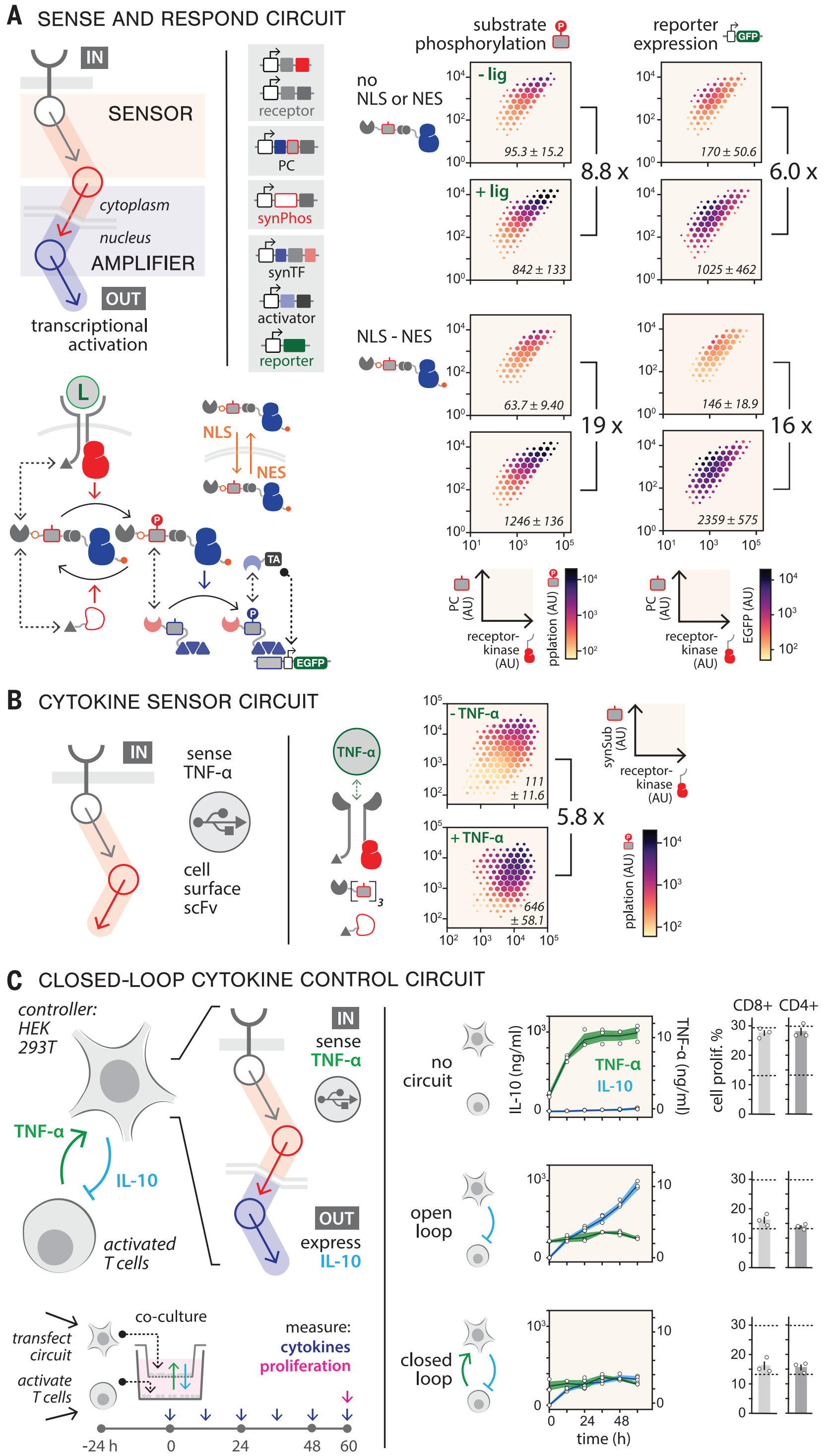


**Fig. 3. Constructing receptor-coupled sensor circuits.** (**A**) Sensor circuit design. (Left) Reversible ligand binding to extracellular receptors triggers phosphorylation cycle activation. The circuit is encoded as four genes on three plasmids; it includes two receptor chains (FRB*-TM-kinase domain and FKBP-TM-LZ), a synSub with three motif repeats, and a synPhos. (Right) Dimerization of extracellular FRB* and FKBP domains induced by ligand (AP21967) triggers colocalization of receptor-appended kinase with synSub (cyan box), leading to phosphorylation. (**B**) Testing receptor-induced phosphorylation. HHH plots depict flow-cytometry data from sensor circuit compositions analyzed 12 hours after treatment with 200 nM ligand AP21967 (+ lig) or a carrier-only control (– lig). Values in each plot indicate mean phosphorylation (AU) ± SEM (*n* = 3 biological replicates). Phosphorylation fold-change values are next to each set of plots. (Bottom right) Dose response plot shows synSub phosphorylation as a function of ligand concentration. Individual triplicate measurements (circles) with mean ± SEM (error bars) are shown. Data were fitted (magenta line) to eq. S150 (supplementary materials) ($n_H$, Hill coefficient; $EC_{50}$, input at 50% activation). (**C**) Measuring pathway activation by using a LLP condensate colocalization reporter. (Left) The reporter consists of EGFP-tagged synSub and an mCherry-tagged SH2 domain tagged with an intrinsically disordered protein (IDP); phosphorylation leads synSub-EGFP recruitment to condensates and EGFP-mCherry colocalization. For activation experiments, cells were cultured for 24 hours, and then time-lapse images were taken every 10 min after ligand addition over an 80-min time course to track pathway activation (left time course plot). For deactivation experiments, 200 nM ligand was added for 90 min followed by addition of the inhibitor imatinib mesylate (10 $\mu$M), and then time-lapse images were taken every 5 min after inhibitor addition over a 60-min time course (right time course plot). Data are plotted as single-cell trajectories (thin pink lines) for activation (*n* = 15 cells) and deactivation (*n* = 10 cells), with mean values (thick pink line) ± SEM (shaded pink bands). (Middle right) For selected time points (0, 30, and 70 min for activation and 0 and 30 min for deactivation), images of EGFP and mCherry (false-colored green and red, respectively) are shown for representative single cells, with the cell boundaries (dotted white outlines) as determined with custom segmentation software. (Top right) Histograms show max-normalized EGFP and mCherry intensities along the straight white dashed lines drawn in images, and intensity is plotted for each channel in the same plot, with the black dashed line indicating max-normalized cytoplasmic EGFP intensity; shaded regions indicate outside the cytoplasm. Scale bars, 5 $\mu$m.

engineered a reporter system that allowed us to track in real time with fluorescence microscopy the accumulation of phosphorylated synSub (Fig. 3C). The reporter was created by fusing an SH2 domain and mCherry fluorescent protein to a tag derived from polar organizing protein Z (*41*), an intrinsically disordered protein that can sequester fused client proteins into cytoplasmically localized liquid-liquid phase (LLP) condensates (fig. S21A) (*42*). This enabled us to monitor circuit activation by quantifying the colocalization of EGFP-fused synSub to the condensates as a proxy for phosphorylation (fig. S21, B and C, and movie S1). We detected EGFP-mCherry colocalization within 10 min after ligand addition, and steady state was reached after ~1 hour, whereas addition of a synKin inhibitor to the fully active

**Fig. 4. Phosphorylation circuit-mediated closed-loop sense-and-response function.** (**A**) Using phospho-signaling to connect extracellular sensing to transcriptional output. (Left) The phospho-sensor circuit (Fig. 3) was coupled to a two-step amplifier circuit (Fig. 2), which was in turn coupled to a phosphorylation-dependent transcriptional reporter module (fig. S19), yielding a membrane-to-nucleus signaling pathway. The PC is tagged with nuclear localization (NLS) and nuclear export signals (NES) to enable shuttling between the nucleus and cytoplasm. (Right) HHH plots for (left column) PC phosphorylation and (right column) EGFP expression are shown for cells ± ligand for compositions with and without NLS or NES tags. Values in the plots indicate mean fluorescence (AU) ± SEM (*n* = 3 biological replicates). Fold-change values are shown at right of each set of plots. (**B**) Cytokine sensor circuit. Receptor chains in Fig. 3 are appended to (single-chain variable fragments) scFvs that bind to human TNF-α. TNF-α binding induced synSub phosphorylation. (Right) HHH plots depict flow cytometry data analyzed 12 hours after treatment with 20 ng/ml TNF-α (+ lig) or a carrier-only control (− lig). Values in each plot indicate mean phosphorylation (AU) ± SEM (*n* = 3 biological replicates). (**C**) Engineering a phospho-signaling pathway for closed-loop therapeutic control. (Top left) HEK293T cells expressing a circuit that can sense TNF-α (green arrow) and respond by secreting IL-10 (blue arrow) are (bottom left) placed in transwell coculture with activated T cells for 60 hours, with media collected every 12 hours to measure cytokine levels. T cell proliferation was assessed by means of EdU assay at 60 hours. TNF-α and IL-10 time courses are shown for empty HEK293T cells (no circuit), constitutive IL-10 expression driven by a nonreceptor synKin driving phosphorylation of 2-step cascade (open loop), and the sense-and-respond circuit (closed loop). (Middle) Each circle indicates a different PBMC replicate (black line, mean values; shaded regions, ±SEM). Measurements of $CD4^+$ and $CD8^+$ proliferation, with circles indicating data from three PBMC biological replicates. Error bars indicate mean values ± SEM [*n* = 3 measurements; top dashed line, proliferation (maximum EdU signal) of activated T cells alone with no HEK293T cell; bottom dashed line, proliferation of activated T cells inhibited with 500 ng/ml IL-10 (minimum EdU signal)].



pathway led to rapid synPhos-dependent delocalization (Fig. 3C and movie S2). In a circuit in which synPhos was absent, delocalization occurred >10× slower, demonstrating that synPhos is critical for rapid circuit reversibility (fig. S22A). Fitting these data to a dynamic model (fig. S22B) yielded circuit activation and deactivation half-times of 28.8 and 22.8 min, respectively—rapid dynamics similar to those measured for JAK/STAT (Janus kinase/signal transducer and activator of transcription) (*43*) and TGF-β (transforming growth factor–β) cytokine signaling pathway (fig. S22C) (*44*).

We tested whether we could combine our phospho-sensor and two-step amplifier modules to create a sense-and-respond circuit that would convert an extracellular input signal into expression of a transgene (Fig. 4A, left, and fig. S23A). To promote membrane-to-nucleus signal propagation, we selected an amplifier composition [Fig. 2C, (2)] that showed both high gain and high downstream phosphorylation, and we appended nuclear localization sequence (NLS) and nuclear exporting sequence (NES) to the PC to promote shuttling between the cytoplasm and nucleus. Regulation of transcriptional activation was implemented by fusing the second substrate to a synthetic zinc finger transcription factor (synTF) (*10*) to facilitate phospho-dependent recruitment of an SH2-fused transcriptional activation domain (TAD), resulting in initiation of EGFP reporter expression. We tested this six-protein, seven-gene circuit in HEK293T cells and observed PC phosphorylation (19× fold-change) and EGFP expression (16× fold-change) in response to ligand addition (Fig. 4A, right). Circuit designs not tagged with a NLS or NES (fig. S23B) and lower-gain circuit designs (fig. S24A) showed <10× fold-change, underscoring the importance of shuttling and amplification as circuit design features. The circuit also demonstrated a linear dose-response profile ($n_H$ = 0.97) (fig. S24B).

As a demonstration of the translational potential of our framework, we engineered a circuit that senses tumor necrosis factor–α (TNF-α)—a cytokine secreted by T cells that drives adverse inflammatory response—and responds by secreting interleukin 10 (IL-10), a cytokine that inhibits T cell activation, proliferation, and TNF-α production but has toxic side effects that limit its clinical utility (*45*). This circuit could establish an anti-inflammatory control loop that suppresses T cell activation while maintaining low concentrations of both cytokines. We tested this by reconfiguring the sense-and-respond circuit shown in Fig. 4A; we appended single-chain antibody fragments (scFvs) that recognize TNF-α to the receptors (Fig. 4B and fig. S25) and replaced the EGFP reporter with IL-10. HEK293T cells containing this circuit were introduced into a transwell coculture with CD3- and CD28-activated human peripheral blood mononuclear cells (PBMCs), and cytokine production and T cell proliferation were assessed across a 60-hour time course (Fig. 4C, left). Cocultures containing cells with no circuit showed rapid accumulation of TNF-α and robust T cell proliferation, whereas cells containing a circuit driving constitutive IL-10 expression (open-loop composition) inhibited TNF-α secretion and T cell proliferation (Fig. 4C, right, and fig. S26). Cells containing the sense-and-respond circuit (closed-loop composition) also suppressed T cell proliferation but reached low steady-state concentrations of both TNF-α and IL-10 after ~12 hours. As indicated by modeling the dynamics of this system (fig. S27), this rapid setpoint convergence is likely dependent on the fast activation and deactivation rates of our phospho-signaling circuit and may not be achievable with circuits that use molecular mechanisms that operate on slower timescales (figs. S27 and S28).

We engineered synthetic phospho-signaling circuits using a simple design logic in which phosphorylation cycles are used as building blocks, and circuit connectivity and information flow are defined through programmed protein-protein interactions. Our design framework enabled predictive tuning of circuit behavior and the use of nonequilibrium thermodynamic modeling to guide circuit design. Although the part set we used in this study to demonstrate the practicability of our design scheme consisted largely of domains and motifs repurposed from native human immune signaling, our framework should facilitate incorporation of domains drawn from other sources or generated through computational design (*46*). Catalytic domains could be engineered to enhance circuit performance through activity tuning or by introducing allosteric regulation. Because the functional specificities of our components are determined by recruitment, scaling to greater circuit complexity could be enabled by simply expanding the number of orthogonal interaction domains in our part set. This could facilitate the construction of circuit topologies that carry out advanced signal processing functions, such as Boolean logic enabled through multisite phosphorylation, feedback connections that tune circuit dynamics or introduce ultrasensitivity (*47*), or multi-input-output circuits that can perceive and compute internal or external states (*48*).

Last, because our circuits are posttranslational and signal rapidly and reversibly, they can potentially support a broad array of cell-based diagnostic and therapeutic applications that require sensing of minute-scale physiological or pathological events (*4*). The plug-and-play configurability of our circuits should enable their coupling to diverse receptor inputs capable of sensing small molecules, bioactive factors, or disease markers (*49*). Because of their temporal responsiveness, the circuits may complement or offer advantages over other highly programmable circuit design schemes that signal through slower-turnover molecular mechanisms (such as transcription or proteolysis) (*50*). Additionally, because our circuits operate in parallel to native signaling pathways, they offer opportunities for programming signal-processing functions that are not possible for signaling circuits that harness native components to propagate signal. Furthermore, because they can be configured with human-derived protein domains and are relatively compact, circuits constructed by using our design framework are likely to have low immunogenicity (fig. S29) and could potentially be delivered to clinically relevant cell types (fig. S30) to enable therapeutic sense-and-respond function.

## REFERENCES AND NOTES

1. T. Pawson, J. D. Scott, *Trends Biochem. Sci.* **30**, 286–290 (2005).
2. W. A. Lim, *Nat. Rev. Mol. Cell Biol.* **11**, 393–403 (2010).
3. R. Kojima, D. Aubel, M. Fussenegger, *Cell. Mol. Life Sci.* **77**, 3567–3581 (2020).
4. C. J. Bashor, I. B. Hilton, H. Bandukwala, D. M. Smith, O. Veiseh, *Nat. Rev. Drug Discov.* **21**, 655–675 (2022).
5. D. E. Cameron, C. J. Bashor, J. J. Collins, *Nat. Rev. Microbiol.* **12**, 381–390 (2014).
6. C. J. Bashor, J. J. Collins, *Annu. Rev. Biophys.* **47**, 399–423 (2018).
7. A. S. Khalil *et al.*, *Cell* **150**, 647–658 (2012).
8. C. A. Voigt, *Curr. Opin. Biotechnol.* **17**, 548–557 (2006).
9. A. A. Nielsen *et al.*, *Science* **352**, aac7341 (2016).
10. C. J. Bashor *et al.*, *Science* **364**, 593–597 (2019).
11. J. M. Skerker *et al.*, *Cell* **133**, 1043–1054 (2008).
12. W. R. Whitaker, S. A. Davis, A. P. Arkin, J. E. Dueber, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18090–18095 (2012).
13. S. R. Schmidl *et al.*, *Nat. Chem. Biol.* **15**, 690–698 (2019).
14. A. Mazé, Y. Benenson, *Nat. Chem. Biol.* **16**, 179–187 (2020).
15. L. Scheller *et al.*, *Nat. Commun.* **11**, 3085 (2020).
16. R. D. Jones *et al.*, *Nat. Commun.* **13**, 1720 (2022).
17. D. Mishra, P. M. Rivera, A. Lin, D. Del Vecchio, R. Weiss, *Nat. Biotechnol.* **32**, 1268–1275 (2014).
18. D. Mishra *et al.*, *Science* **373**, eaav0780 (2021).
19. R. M. Gordley *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13528–13533 (2016).
20. H. K. Chung *et al.*, *Science* **364**, eaat6982 (2019).
21. L. Schukur, B. Geering, G. Charpin-El Hamri, M. Fussenegger, *Sci. Transl. Med.* **7**, 318ra201 (2015).
22. K. Krawczyk, L. Scheller, H. Kim, M. Fussenegger, *Nat. Commun.* **11**, 608 (2020).
23. A. Goldbeter, D. E. Koshland Jr., *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6840–6844 (1981).
24. D. Bray, *Nature* **376**, 307–312 (1995).
25. T. Pawson, P. Nash, *Genes Dev.* **14**, 1027–1047 (2000).
26. R. P. Bhattacharyya, A. Reményi, B. J. Yeh, W. A. Lim, *Annu. Rev. Biochem.* **75**, 655–680 (2006).
27. Z. Chen, M. B. Elowitz, *Cell* **184**, 2284–2301 (2021).
28. S. H. Park, A. Zarrinpar, W. A. Lim, *Science* **299**, 1061–1064 (2003).
29. C. J. Bashor, N. C. Helman, S. Yan, W. A. Lim, *Science* **319**, 1539–1543 (2008).
30. A. H. Courtney, W. L. Lo, A. Weiss, *Trends Biochem. Sci.* **43**, 108–123 (2018).
31. A. Acharya, S. B. Ruvinov, J. Gal, J. R. Moll, C. Vinson, *Biochemistry* **41**, 14122–14131 (2002).
32. D. M. Underhill, H. S. Goodridge, *Trends Immunol.* **28**, 66–73 (2007).
33. Y. Cui, G. Sun, *Protein Sci.* **28**, 533–542 (2019).
34. J. P. Ferreira, K. W. Overton, C. L. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11284–11289 (2013).
35. L. Tautz, D. A. Critton, S. Grotegut, *Methods Mol. Biol.* **1053**, 179–221 (2013).
36. J. A. Ubersax, J. E. Ferrell Jr., *Nat. Rev. Mol. Cell Biol.* **8**, 530–541 (2007).
37. T. Pawson, *Cell* **116**, 191–203 (2004).
38. E. A. Ottinger, M. C. Botfield, S. E. Shoelson, *J. Biol. Chem.* **273**, 729–735 (1998).
39. P. B. Detwiler, S. Ramanathan, A. Sengupta, B. I. Shraiman, *Biophys. J.* **79**, 2801–2817 (2000).
40. B. Z. Stanton, E. J. Chory, G. R. Crabtree, *Science* **359**, eaao5902 (2018).
41. G. Ebersbach, A. Briegel, G. J. Jensen, C. Jacobs-Wagner, *Cell* **134**, 956–968 (2008).
42. S. Boeynaems *et al.*, *Trends Cell Biol.* **28**, 420–435 (2018).
43. H. Kiu, S. E. Nicholson, *Growth Factors* **30**, 88–106 (2012).
44. B. Labibi, M. Bashkurov, J. L. Wrana, L. Attisano, *iScience* **23**, 101416 (2020).
45. A. Saxena *et al.*, *Cytokine* **74**, 27–34 (2015).
46. N. B. Woodall *et al.*, *Nat. Struct. Mol. Biol.* **28**, 762–770 (2021).
47. J. E. Ferrell Jr., S. H. Ha, *Trends Biochem. Sci.* **39**, 612–618 (2014).
48. Z. Chen *et al.*, *Science* **386**, 1243–1250 (2024).
49. J. Manhas, H. I. Edelstein, J. N. Leonard, L. Morsut, *Nat. Chem. Biol.* **18**, 244–255 (2022).
50. X. J. Gao, L. S. Chong, M. S. Kim, M. B. Elowitz, *Science* **361**, 1252–1258 (2018).

## ACKNOWLEDGMENTS

We thank members of the Bashor laboratory for helpful discussions. We acknowledge C.-H. Wang for his contributions to development of the modeling framework. We also thank J. Toettcher and S. Boeynaems for advice on engineering and monitoring protein condensates. We thank O. Veiseh and S. Fleury for providing APRE-19 cells. We thank Y. Wang for helping with immunogenicity analysis, R. O'Connell for helping with APRE-19 transfection, and P. Pungchai for assisting with viability analysis. We also thank R. Butcher for setting up the Nikon NASPARC imaging system for taking time-lapse movies. We also thank S. Li and J. Her for helping with cloning. We also thank G. Bedford, J. Li, and I. Hilton for setting up immunoblotting assays.

**Funding:** This work was supported by grants from NIH R01 EB029483 (C.J.B.), NIH R01 EB032272 (C.J.B.), NIH R21 NS116302 (S.D.O. and C.J.B), ONR N00014-21-1-4006 (C.J.B.), NIH NIGMS 5R35GM119461 (J.W.R. and P.M.), the Robert J. Kleberg Jr. and Helen C. Kleberg Foundation (C.J.B.), the Claire Glassell Pediatric Fund (S.D.O.), the Grace Reynolds Wall Research Fund (S.D.O.), and NSF GRFP award 1842494 (A.J.W.). **Author contributions:** C.J.B., N.M.D., and X.Y. conceived of the study. C.J.B. and N.M.D. carried out initial experiments. X.Y. built all DNA constructs and carried out experiments to generate the data published herein, with assistance from K.J., A.J.W., J.L., and J.N.; J.W.R., X.Y., K.R., and K.J. developed the quantitative modeling framework, with input and supervision from P.M. and C.J.B.; X.Y. and A.J.W. developed the transwell assay, with supervision from S.D.O. and C.J.B.; K.R. developed image analysis and dynamic circuit modeling. X.Y., K.R., J.W.R., P.M., and C.J.B. analyzed the data. C.J.B., N.D., and J.J.C supervised the study. X.Y. and C.J.B. wrote the manuscript, with generous input from all authors. **Competing interests:** A provisional patent application that covers technologies described in this manuscript has been filed by Rice University. **Data and code availability:** All data are available in the manuscript or supplementary materials. Sequences of all DNA constructs (table S1) are uploaded to the following Github repository and will be uploaded to Addgene later. All computational data generated in this study and modeling or analysis software code are available at https://github.com/Emergent-Behaviors-in-Biology/syn-phospho-circuits. Correspondence and request for materials should be addressed to C.J.B.

**SUPPLEMENTARY MATERIALS**

science.org/doi/10.1126/science.adm8485
Materials and Methods
Supplementary Text
Figs. S1 to S30
Tables S1 to S6
References (*51–157*)
Movies S1 and S2

Submitted 9 November 2023; accepted 24 October 2024
10.1126/science.adm8485

---

EVOLUTIONARY BIOLOGY

# Competitive social feedback amplifies the role of early life contingency in male mice

**Matthew N. Zipple[1]*, Daniel Chang Kuo[1]†, Xinmiao Meng[1]‡, Tess M. Reichard[1], Kwynn Guess[1], Caleb C. Vogt[1], Andrew H. Moeller[1,2], Michael J. Sheehan[1]***

**Contingency (or "luck") in early life plays an important role in shaping individuals' development. By comparing the developmental trajectories of functionally genetically identical free-living mice who either experienced high levels of resource competition (males) or did not (females), we show that competition magnifies early contingency. Male resource competition results in a feedback loop that magnifies the importance of early contingency and pushes individuals onto divergent, self-reinforcing life trajectories, while the same process appears absent in females. Our results indicate that the strength of sexual selection may be self-limiting, and they highlight the potential for contingency to lead to differences in life outcomes, even in the absence of any underlying differences in ability ("merit").**

Contingency (colloquially, "luck" or "chance") has long been recognized as an important determinant of outcomes in both biological and social sciences (*1–15*). The contingency hypothesis posits that an individual's behavior, health, social position, and fitness are strongly dependent on events and experiences that it can neither control nor predict (*5*, *16–20*). The outcomes of contingent events in early life are often especially important, as they can set individuals onto divergent, self-reinforcing trajectories (*1*, *2*, *5*, *17*, *18*, *20*). Recent evolutionary theory has argued that luck in an individual's life, particularly in early life, can outweigh individual quality in determining lifetime reproductive success (*1*, *2*).

Many animals naturally live within larger social groups, such that contingency in event outcomes is inextricably tied to individuals' relationship to the behavior of others within societies (*5*, *21–25*). Through repeated social interactions, individuals adopt a consistent set of social phenotypes [i.e., their "social niche" (*22*, *24*, *26*)]. We hypothesize that competitive social processes magnify the importance of contingency in early life. For example, animals that begin with zero or small differences in competitive ability may differ in their access to resources owing to variation in contingent dominance or territorial interactions (*19*, *27–30*). The resulting increased resource access for a subset of the population then improves those animals' condition relative to those with reduced resource access, further entrenching the initial differences and magnifying the importance of early contingency (*19*, *21*, *30–33*). This process is analogous to the "Matthew effect" in the social sciences, a phenomenon in which individuals or institutions that achieve early success tend to achieve ever greater success in the future (*34–36*).

Experimentally studying the role of contingency in individual outcomes is achievable with the use of "replicate individuals" that allow researchers to effectively "replay the tape of life" for a single genotype under different circumstances (*37*, *38*). Within the lab, studies of functionally genetically identical animals indicate that contingent microenvironmental differences during development cause between-individual differences in early behavior, which then increase in magnitude over time (*39–45*). Yet assessing the ways in which competitive social processes interact with early contingency requires the study of replicate individuals living under realistic, complex, dynamic social conditions—requirements that cannot be readily met under standardized laboratory conditions (*46–48*). We overcame this limitation by studying the development of individuality in ecologically relevant spatial and social behaviors from infancy through adulthood in age-matched, isogenic (hereafter, "genetically identical") mice living outside in a shared, seminatural field environment (figs. S1 and S2).

## Males develop individual adult behavioral phenotypes earlier than females

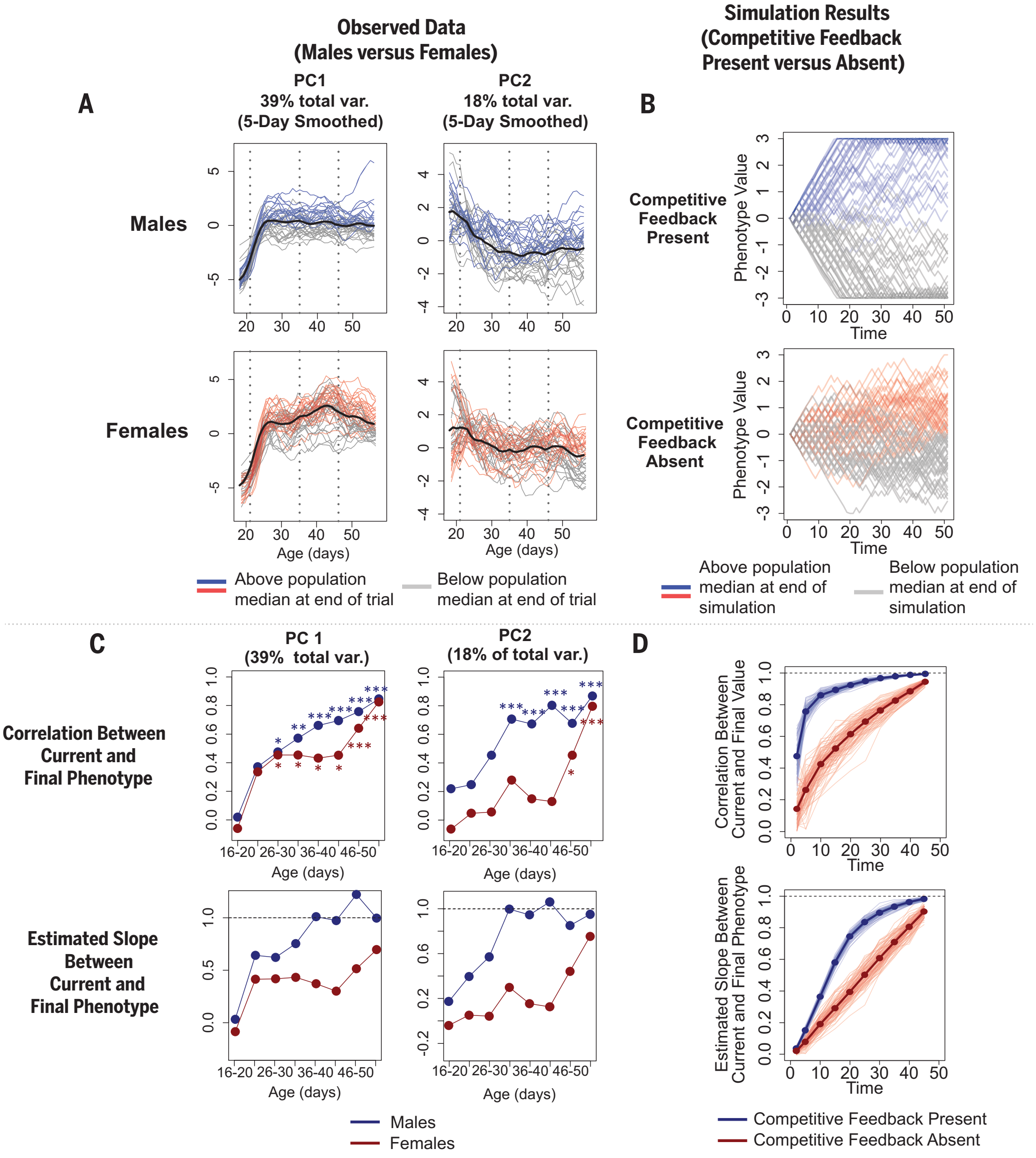
In free-living C57BL/6J lab mice, males compete for territorial control and resource access, while females do not (*49*, *50*). Females appear unconstrained in their movement, whereas males repeatedly visit only a small subset of available space (*49*, *50*) (fig. S2). Males also spend much less time engaged in spatiotemporal overlap with each other than do females (*49*, *50*) (fig. S2). And territory-less males have higher mortality and attain less access to females than do territory holders (*49*, *50*) (fig. S2). We hypothesized that this difference in competitive experiences causes males to diverge onto self-reinforcing developmental trajectories as some win competitive interactions and others lose. We expected this same dynamic to be absent in females.

In this study, we monitored 16 litters of the C57BL/6J inbred mouse strain ($n$ = 104 pups, 90 of which survived to adulthood) from infancy through adulthood outside in a large (~560 m$^2$) enclosure that emulated the natural foraging and social environment of commensal house mice (fig. S1, A to C). We placed 2-week-old litters and their mothers in the enclosures within one of 16 identical "resource zones" containing food and shelter, monitored with radio-frequency identification (RFID) antennae. Our goal was to minimize the impacts of genetic variation, maternal effects, and physical microenvironmental differences, thereby restricting environmental variation to social dynamics as much as possible. We inferred periods of social overlap using an established workflow to translate RFID positional data into estimates of the duration of social aggregations within each of the monitored zones (fig. S1D)

---

[1]Department of Neurobiology and Behavior, Cornell University, Ithaca, NY, USA. [2]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA.
***Corresponding author. Email: mz343@cornell.edu (M.N.Z.); msheehan@cornell.edu (M.J.S.)**
†Present address: Department of Integrative Biology, University of California, Berkeley, Berkeley, CA, USA.
‡Present address: School of Public Health, Yale University, New Haven, CT, USA.

**Fig. 1. Males adopt their adult phenotypes earlier than females.** (**A**) Traces of observed individual behavioral PC1 and PC2 values, smoothed over 5 days, across animals' development. (**B**) Traces of the development of individual differentiation from a single run of agent-based simulations in which individuals' phenotypes develop either in the presence or absence of competitive feedback mechanisms. (**C** and **D**) Quantification of the patterns in (A) and (B). (C) (Top row) The correlation between earlier and final adult behavior (*y* axis) is stronger in males for PC1 (left column) and PC2 (right column). Asterisks denote significance of the correlations depicted in each point (linear model, Bonferroni corrected for 16 comparisons in each sex, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$). (C) (Bottom row) The slope of the relationship between earlier and final adult behavior (*y* axis). The slope of this relationship is consistently closer to 1 for males than for females. (D) Results from 1000 iterations of the agent-based simulation.



(*49*). On the basis of 7.4 million RFID reads, we traced the development of 17 social and spatial phenotypes from infancy through adulthood (14 to 58 days; table S1).

We first showed that genetically identical animals display distinct, individually repeatable social behaviors under ecologically relevant contexts, an open question in behavioral ecology (*37*, *38*, *44*, *45*, *51*, *52*). We measured repeatability as the proportion of a phenotype's total variation in each sex that was explained by individual identity (*53*) over a sliding 5-day age window, after controlling for maternal and litter identity. We detected significant repeatability across 5-day periods in all measured phenotypes, emerging as early as age 21 days (fig. S3), roughly 1 month earlier than reported for spatial behavior of female populations of this strain in enriched lab vivaria (*11*, *27*). Although essentially all phenotypes that we measured became highly repeatable in both sexes over these short time frames, many more phenotypes became significantly predictive over longer time periods (15 to 25 days) in males than in females (fig. S3I).

We next assessed the developmental timing at which males and females assumed their individually distinct adult behavioral phenotypes. Many of the phenotypes that we measured covaried, so interpreting our data required us to flatten our phenotypes into orthogonal measures of behavior. We used principal components analysis to reduce the dimensionality of 16 of our 17 behavioral phenotypes into two principal components (PCs ) that accounted for a majority of the total variation in our dataset (57% across PC1 and PC2; "time of first nightly transition" values were missing in young animals who did not move between zones). PC1 is a course metric of social and spatial exploration, being inversely related to both spatial and social fidelity. PC2 is a metric of social connectivity, being positively correlated with all prosocial traits, including number of social partners and the strength of social connections (table S2).

We identified animals' final adult behavioral phenotypes by taking the average of each individual's PC1 and PC2 scores during the last 3 days of the experiment (age 56 to 58 days; Fig. 1A). For each sex, we then assessed the relationship between individuals' phenotypes at earlier time points and these final adult behavioral phenotypes by building linear regressions between final adult phenotype and individuals' average phenotypes over 5-day, nonoverlapping windows (e.g., age 21 to 25 days).

For both PC1 and PC2, males assumed their individual adult behavioral phenotypes earlier than females, consistent with our analysis of single behaviors (Fig. 1C and fig. S3I). Male behavior became predictive ($P < 0.05$, Bonferroni corrected for 16 hypothesis tests) of final

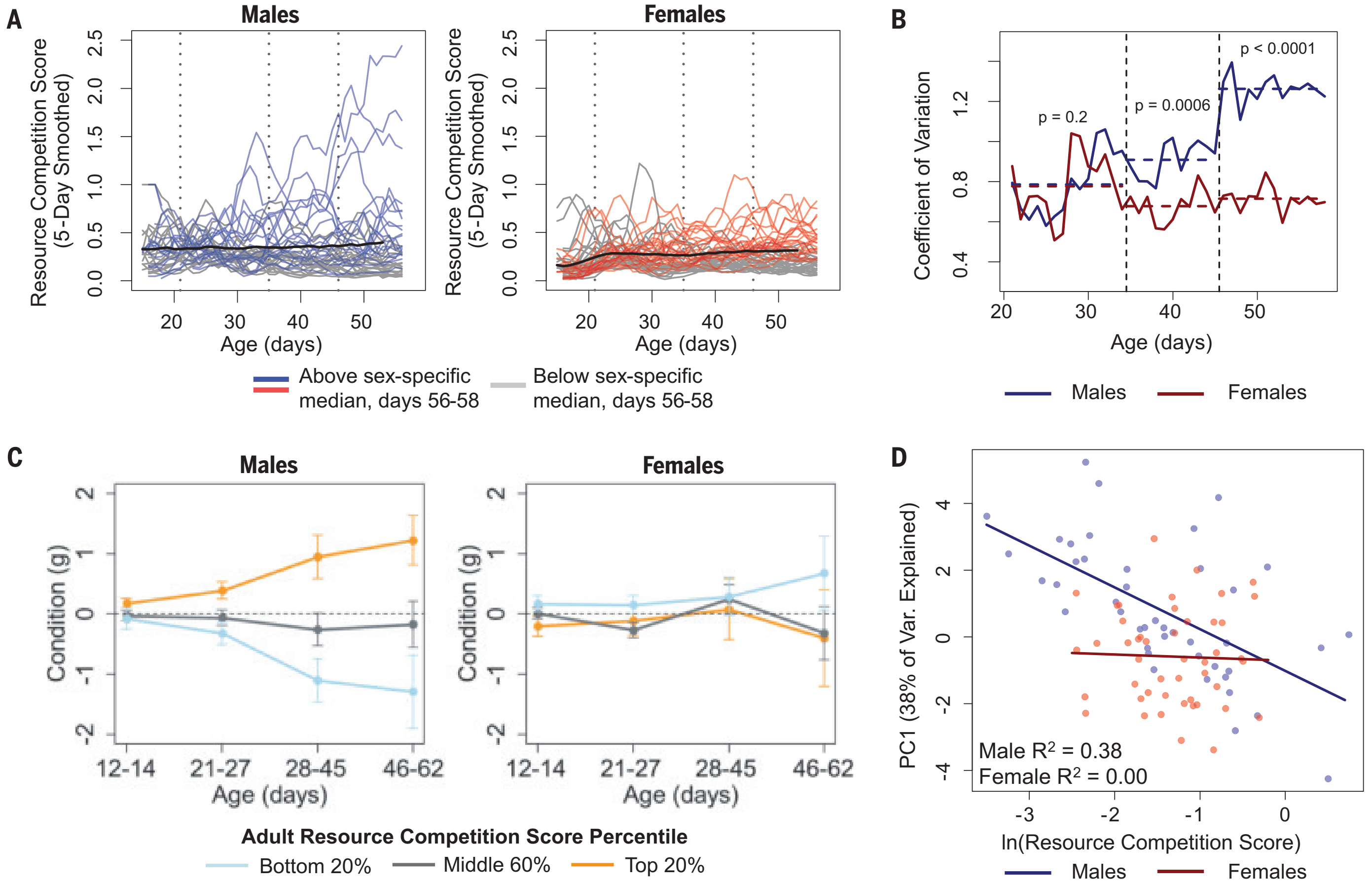


**Fig. 2. Competition for resource access shapes males' phenotypes through a sex-specific competitive feedback loop.** (**A**) Traces of individual resource competition scores, smoothed over 5 days. Black lines indicate sex-specific means. Vertical dotted lines indicate approximate ages of weaning ("juvenility," 21 days), sexual maturation ("adolescence," 35 days), and onset of conceptive mating ("adulthood," 46 days). (**B**) Male resource competition scores are more variable across individuals than female scores (see also fig. S2). Horizontal dashed line segments indicate the average coefficient of variation across each developmental stage. (**C**) Male adult resource competition scores are predicted by small differences in body mass in early life, a difference that is magnified over time. The relationship is absent in females. The *y* axes represent deviations from age-predicted body mass, points represent averages ± SEM. (**D**) Adult resource competition score strongly predicts an integrative measure (PC1) of the 16 other spatial and social phenotypes in males but not in females.

adult behavior earlier than did female behavior (PC1: 26 days versus 26 days, PC2: 31 days versus 46 days; Fig. 1C). The strength of the correlation between earlier and later behavior for the same individual was substantially higher for males than for females across development (PC1: days 31 to 50; PC2: days 16 to 50; Fig. 1C). Moreover, the slope of the linear relationship between earlier behavior and adult behavior was closer to 1 for males than for females, and this strong relationship developed at an earlier age for both PC1 and PC2 (Fig. 1C and fig. S4). The above results were unchanged when we generated separate sex-specific PC values instead of including both sexes in a single analysis (fig. S5).

These differences in the developmental timing of behavioral individuality are unlikely to be explained by sex differences in the timing of sexual maturation, as males and females of this strain display similar timings of the onset of puberty and successful mating (*54*–*59*). Differences in final adult values of PC1 and PC2 were also not predicted in either sex by maternal identity or the area of the enclosure in which animals started the experiment (table S3).

We next assessed whether these sex differences in the developmental timing of individuality could be explained by sex-specific differences in the importance of competitive feedback in amplifying the impact of early contingency. We built a quantitative agent-based model to generate predictions of how competitive processes shape the long-term phenotypic impacts of early contingency (Fig. 1, B and D). We assumed that all individuals began with the same value of a phenotype, which then changed at discrete timesteps across their lives. Without competitive feedback, phenotype values increased or decreased randomly. With competitive feedback present, an individual's phenotype increased in value if it won a competitive interaction with a randomly chosen individual and declined if it lost. The probability of winning the interaction depended on the relative phenotypic values of the two interactants. With competition present, event outcomes are therefore least predictable at the beginning of life and become more predictable as phenotypic differences emerge and are reinforced through competitive feedback (see supplementary materials).

The qualitative results of the simulation closely mirrored our observations of differences in the development of individuality in males and females in our system (compare Fig. 1A to Fig. 1B and Fig. 1C to Fig. 1D). When competitive feedback loops are present, (i) the correlation between behavior at any given time and behavior at the end of the modeled period is stronger, and (ii) the slope of the relationship between earlier and later behavior is closer to 1.0 (Fig. 1, C and D). Thus, the sex difference that we observed in the development of behavioral individuality could be entirely explained by differences in sex-specific competitive processes that amplify contingent early life differences in phenotype.

### Competition for resource access acts as a sex-specific competitive feedback loop

We next assessed whether males and females displayed differences in the strength of resource competition in a fashion that would support this putative sex-biased competitive feedback loop. We estimated individuals' nightly resource access by calculating a nightly "resource competition score" for each animal (see methods in the supplementary materials). Consistent with males facing higher levels of competition than females, the resource competition score varied more among males than it did among females (Fig. 2A). This difference emerged concurrently with the onset of sexual maturity, the period when we expect intrasexual competition to increase in intensity among male mice (fig. S2C).

Two additional pieces of evidence are consistent with males, but not females, experiencing strong competitive feedback that set them on self-reinforcing divergent life trajectories. First, small individual differences in early body mass (days 21 to 27) predicted adult (days 46 to 58) resource competition scores for males ($P < 0.05$; Fig. 2C and fig. S6) but not for females. Consistent with individuals starting out on an approximately "even playing field," very minor differences in infant body condition before release into the field (days 12 to 14) did not predict future resource access in either sex ($P > 0.05$; fig. S6). Differences in body mass between males with differential resource access then increased over the course of development (age × adult resource competition score interaction: $P = 0.004$; Fig. 2C and fig. S6), consistent with males experiencing a competitive feedback loop that increased the condition of winners relative to losers. This developmental pattern was absent in females ($P = 0.19$; Fig. 2C).

Second, male resource competition scores strongly predicted the other male behavioral phenotypes, whereas the same was not true of females, indicating that competition in males has major impacts on males' daily behaviors and access to mates. Specifically, we performed a second principal components analysis using the 16 phenotypes other than resource competition score that we measured in adulthood (age 46 to 58 days) to obtain a single integrative measure of animals' other adult behavior (PC1 explained 38% of the total variation in this dataset). Males' average adult resource competition scores strongly predicted this adult PC1 value [$P < 0.0001$, coefficient of determination ($R^2$) = 0.38; Fig. 2D], whereas females' scores did not ($P = 0.86$, $R^2 = 0.00$; Fig. 2D). The same conclusion holds if we assess individual adult phenotypes, including access to mates (table S4).

### Discussion

Our results provide empirical support for the hypothesis that contingency (or "luck") in early life can have a major and competition-dependent impact on the development of animals' individual differences in social and fitness-relevant phenotypes. The outcomes we observed are not explained by genotype or macroenvironment (which we controlled), maternal identity, or initial location in the arena. To the extent that early success or failure in social competition was determined by individual quality, the factors that generated individual differences were contingent experiences. Competitive feedback loops magnify the importance of contingency experienced early in life, such that young free-living male lab mice enter onto divergent, self-reinforcing developmental trajectories. Our results match an agent-based simulation of expected differences in the developmental timing of individuality in the presence and absence of competitive feedback.

We expect the sex-specificity and intensity of competitive feedback loops to vary across different contexts and species, depending on social behavioral ecology. Here, the most relevant differences in social competition happen to be related to sex, but this need not be the case [e.g., in wild house mice, females compete for space and mates under some conditions (*60*, *61*)]. In female-dominant species, the reverse pattern may be present, such that females' outcomes would be more dependent on early luck than males' outcomes (*62*, *63*). In other settings, perhaps including some or many human populations, both males and females may experience high or low levels of competition, or this competition may vary across time or space.

Our results suggest an inevitable limitation of sexual selection's ability to shape behaviors. Intrasexual selection relies on within-sex competition resulting in differential reproductive success, and for variation in this success to be heritable (*64*–*66*). But here we have shown that intrasexual competition also magnifies the importance of early contingency in later life outcomes in the sex expressing that competition. As the importance of luck in determining individual phenotypic outcomes increases, selection's ability to cause evolution declines (*1*, *2*). Thus, intrasexual selection may be self-limiting, as an increase in the importance of competition in a single sex leads, in turn, to an increase in the importance of contingency in determining individual outcomes in that same sex. This increased importance of luck may help to explain why sexual selection fails to fully deplete genetic variation [i.e., the lek paradox (*67*–*69*)].

Our results provide a strong biological analog to the Matthew effect, an often-observed phenomenon in social science whereby small individual advantages earlier in life are correlated with ever larger advantages over time (*34*–*36*). Such processes are understood to be the result of social feedback mechanisms, by which an individual's initial success improves their opportunities for future success as well as the perception by other members of society of the individual's potential for success (*36*, *70*). Our results suggest that the Matthew effect (i) may have a biological origin, (ii) is especially likely to occur in highly competitive environments or among groups that face high levels of competition, and (iii) may emerge even in the absence of any variation in underlying individual quality or ability. In populations of humans and nonhuman animals, the additional amplifying impacts of competition and contingency exist against the backdrop of unequal starting position and likely magnify early inequalities that result from structural or environmental adversity or advantage (*71*–*74*).

The sources of inequality in human society are of central interest to both moral philosophy and public policy (*75*–*77*). As with reproductive success in nonhuman animals, human outcomes are likely to be partially explained by differences in genotypes (*78*). But we show here that even among isogenic animals, individuals still attain drastically different phenotypic outcomes. Our results add to sociological and biological literature that underscores the potential importance of unpredictable, uncontrollable experiences in generating differences in outcomes even when differences in underlying quality (or "talent") are small or nonexistent (*12*, *15*, *16*).

### REFERENCES AND NOTES

1. R. E. Snyder, S. P. Ellner, G. Hooker, *Am. Nat.* **197**, E110–E128 (2021).
2. R. E. Snyder, S. P. Ellner, *Am. Nat.* **191**, E90–E107 (2018).
3. G. da Col, *Soc. Anal.* **56**, 1–23 (2012).
4. J. B. Losos, *Improbable Destinies: Fate, Chance, and the Future of Evolution* (Riverhead Books, 2017).
5. V. L. Bengtson, G. H. Elder Jr., N. M. Putney, in *Adult Lives: A Life Course Perspective*, J. Katz. S. Peace, S. Spurr, Eds. (Policy Press, 2012), pp. 9–17.
6. S. J. Gould, *Wonderful Life: The Burgess Shale and the Nature of History* (W. W. Norton & Co, 1989).
7. M. Kimura, *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, 1983).
8. R. E. Snyder, S. P. Ellner, *Am. Nat.* **188**, E28–E45 (2016).
9. C. Johnson, *Darwin's Dice: The Idea of Chance in the Thought of Charles Darwin* (Oxford Univ. Press, 2014).
10. N. Rescher, *Luck: The Brilliant Randomness of Everyday Life* (Univ. of Pittsburgh Press, 2001).
11. R. H. Frank, *Success and Luck: Good Fortune and the Myth of Meritocracy* (Princeton Univ. Press, 2016).
12. M. Sauder, *Sociol. Theory* **38**, 193–216 (2020).
13. R. Chetty, N. Hendren, L. F. Katz, *Am. Econ. Rev.* **106**, 855–902 (2016).
14. R. Chetty, N. Hendren, P. Kline, E. Saez, *Q. J. Econ.* **129**, 1553–1623 (2014).
15. M. Gladwell, *The Tipping Point: How Little Things Can Make a Big Difference* (Little, Brown and Company, 2002).
16. R. M. Sapolsky, *Behave: The Biology of Humans at Our Best and Worst* (Penguin Press, 2017).
17. J. Lindström, *Trends Ecol. Evol.* **14**, 343–348 (1999).
18. P. Monaghan, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 1635–1645 (2008).
19. J. E. Smith, B. Natterson-Horowitz, M. M. Mueller, M. E. Alfaro, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **378**, 20220307 (2023).
20. R. Sapolsky, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2401971121 (2024).

21. A. Sih *et al*., *Trends Ecol. Evol.* **30**, 50–60 (2015).
22. K. L. Laskowski, C. C. Chang, K. Sheehy, J. Aguiñaga, *Annu. Rev. Ecol. Evol. Syst.* **53**, 161–182 (2022).
23. J. B. Wolf, E. D. Brodie Iii, A. J. Moore, *Am. Nat.* **153**, 254–266 (1999).
24. R. Bergmüller, M. Taborsky, *Trends Ecol. Evol.* **25**, 504–511 (2010).
25. L. A. Gartland, J. A. Firth, K. L. Laskowski, R. Jeanson, C. C. Ioannou, *Biol. Rev. Camb. Philos. Soc.* **97**, 802–816 (2022).
26. J. B. Saltz, A. P. Geiger, R. Anderson, B. Johnson, R. Marren, *Evol. Ecol.* **30**, 349–364 (2016).
27. R. S. Ostfeld, *Trends Ecol. Evol.* **5**, 411–415 (1990).
28. J. H. Kaufmann, *Biol. Rev. Camb. Philos. Soc.* **58**, 1–20 (1983).
29. T. J. Ord, *Oecologia* **197**, 615–631 (2021).
30. J. E. Smith, B. Natterson-Horowitz, M. E. Alfaro, *Behav. Ecol.* **33**, 1–6 (2022).
31. E. D. Strauss, D. Shizuka, *Proc. Biol. Sci.* **289**, 20220500 (2022).
32. L. A. Dugatkin, R. L. Earley, *Proc. Biol. Sci.* **271**, 1537–1540 (2004).
33. C. Rutte, M. Taborsky, M. W. G. Brinkhof, *Trends Ecol. Evol.* **21**, 16–21 (2006).
34. R. K. Merton, *Science* **159**, 56–63 (1968).
35. T. Bol, M. de Vaan, A. van de Rijt, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4887–4890 (2018).
36. D. Rigney, *The Matthew Effect: How Advantage Begets Further Advantage* (Columbia Univ. Press, 2010).
37. J. Stamps, T. G. G. Groothuis, *Biol. Rev. Camb. Philos. Soc.* **85**, 301–325 (2010).
38. J. A. Stamps, T. G. G. Groothuis, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 4029–4041 (2010).
39. J. Freund *et al*., *Science* **340**, 756–759 (2013).
40. S. Zocher *et al*., *Sci. Adv.* **6**, eabb1478 (2020).
41. G. Kempermann *et al*., *Neurobiol. Dis.* **175**, 105916 (2022).
42. J. S. Kain, C. Stokes, B. L. de Bivort, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 19834–19839 (2012).
43. B. de Bivort *et al*., *Front. Behav. Neurosci.* **16**, 836626 (2022).
44. D. Bierbach, K. L. Laskowski, M. Wolf, *Nat. Commun.* **8**, 15361 (2017).
45. K. L. Laskowski, D. Bierbach, J. W. Jolles, C. Doran, M. Wolf, *Nat. Commun.* **13**, 6419 (2022).
46. G. P. Lahvis, *eLife* **6**, e27438 (2017).
47. G. Lahvis, *Nature* **543**, 623 (2017).
48. M. N. Zipple, C. C. Vogt, M. J. Sheehan, *Neurosci. Biobehav. Rev.* **152**, 105238 (2023).
49. C. C. Vogt *et al*., *BMC Biol.* **22**, 35 (2024).
50. M. N. Zipple, C. C. Vogt, M. J. Sheehan, *Proc. Biol. Sci.* **291**, 20240099 (2024).
51. D. N. Fisher, M. Brachmann, J. B. Burant, *Anim. Behav.* **138**, e1–e6 (2018).
52. F. Trillmich, T. Müller, C. Müller, *Biol. Lett.* **14**, 20170740 (2018).
53. A. M. Bell, S. J. Hankison, K. L. Laskowski, *Anim. Behav.* **77**, 771–783 (2009).
54. B. D'Udine, A. Manning, *Dev. Psychobiol.* **16**, 311–322 (1983).
55. C. Jean-Faucher *et al*., *Int. J. Androl.* **6**, 575–584 (1983).
56. S. K. J. Cross *et al*., *Front. Behav. Neurosci.* **14**, 606788 (2021).
57. R. T. Molenhuis, L. de Visser, H. Bruining, M. J. Kas, *Eur. Neuropsychopharmacol.* **24**, 945–954 (2014).
58. E. A. Witham, J. D. Meadows, S. Shojaei, A. S. Kauffman, P. L. Mellon, *Endocrinology* **153**, 4522–4532 (2012).
59. J. G. Vandenbergh, *Endocrinology* **81**, 345–349 (1967).
60. N. Harrison *et al*., *Front. Zool.* **15**, 4 (2018).
61. A. S. Rusu, S. Krackow, *Behav. Ecol. Sociobiol.* **56**, 298–305 (2004).
62. E. D. Strauss, K. E. Holekamp, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8919–8924 (2019).
63. E. D. Strauss, D. Shizuka, K. E. Holekamp, *Proc. Biol. Sci.* **287**, 20192969 (2020).
64. M. Andersson, *Sexual Selection* (Princeton Univ. Press, 1994).
65. M. West-Eberhard, *Proc. Am. Philos. Soc.* **123**, 222–234 (1979).
66. S. T. Emlen, L. W. Oring, *Science* **197**, 215–223 (1977).
67. J. S. Kotiaho, N. R. LeBas, M. Puurtinen, J. L. Tomkins, *Trends Ecol. Evol.* **23**, 1–3 (2008).
68. J. S. Kotiaho, L. W. Simmons, J. L. Tomkins, *Nature* **410**, 684–686 (2001).
69. A. Pomiankowski, A. P. Møller, *Proc. Biol. Sci.* **260**, 21–29 (1995).
70. P. Hedström, R. Swedberg, Eds., *Social Mechanisms: An Analytical Approach to Social Theory* (Cambridge Univ. Press, 1998).
71. N. Snyder-Mackler *et al*., *Science* **368**, eaax9553 (2020).
72. M. N. Zipple *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2015317118 (2021).
73. S. Nowicki, W. Searcy, S. Peters, *J. Comp. Physiol. A* **188**, 1003–1014 (2002).
74. K. M. E. Wallace, D. W. Hart, F. Venter, A. K. J. van Vuuren, N. C. Bennett, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **378**, 20220310 (2023).
75. J. Rawls, *Justice as Fairness: A Restatement* (Belknap Press, 2001).
76. T. M. Smeeding, *Soc. Sci. Q.* **86**, 955–983 (2005).
77. M. Nico, G. Pollock, Eds., *The Routledge Handbook of Contemporary Inequalities and the Life Course* (Routledge, 2021).
78. K. P. Harden, *The Genetic Lottery: Why DNA Matters for Social Equality* (Princeton Univ. Press, 2021).
79. M. N. Zipple *et al*., Data and Scripts from: Competitive social feedback amplifies the role of early life contingency in male mice, Cornell University eCommons Repository (2024); https://doi.org/10.7298/qcpe-9h62.

## ACKNOWLEDGMENTS

We thank J. Cusker for assistance with animal care. **Funding:** M.N.Z. has been supported by an NSF postdoctoral fellowship in biology (award 2109636) and a Klarman postdoctoral research fellowship from Cornell University. C.C.V. was supported by a Mong Neurotechnology Fellowship from Cornell University. This work was also supported by Pilot and Feasibility awards to M.N.Z. and M.J.S. from the Animal Models for the Social Dimensions of Health and Aging Network (project 5R24AG065172-03). The costs of care for the mouse colony were supported in part by NIGMS award R35GM138284 to A.H.M. A.H.M. was supported by funds from award R35GM138284. **Author contributions:** Conceptualization: M.N.Z. and M.J.S. Methodology: M.N.Z., C.C.V., and M.J.S. Validation: M.N.Z., D.C.K., and C.C.V. Formal analysis: M.N.Z. Investigation: M.N.Z., D.C.K., X.M., T.M.R., K.G., and C.C.V. Resources: M.N.Z., A.H.M., and M.J.S. Data curation: M.N.Z., D.C.K., and C.C.V. Writing – original draft: M.N.Z. and M.J.S. Writing – review & editing: M.N.Z., D.C.K., X.M., T.M.R., K.G., C.C.V., A.H.M., and M.J.S. Visualization: M.N.Z., A.M., and M.J.S. Supervision: M.N.Z. and M.J.S. Project administration: M.N.Z., D.C.K., and M.J.S. Funding acquisition: M.N.Z., A.H.M., and M.J.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data and scripts supporting the analyses in this paper can be found in the Cornell eCommons repository (*79*).

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adq0579
Materials and Methods
Supplementary Text
Figs. S1 to S6
Tables S1 to S4
References (*80–83*)
MDAR Reproducibility Checklist

Submitted 24 April 2024; accepted 8 November 2024
10.1126/science.adq0579

---

**NEURODEVELOPMENT**

# Sensation is dispensable for the maturation of the vestibulo-ocular reflex

**Paige Leary[1]*, Celine Bellegarda[1], Cheryl Quainoo[1], Dena Goldblatt[1,2], Başak Rosti[1], David Schoppik[1]***

**Vertebrates stabilize gaze using a neural circuit that transforms sensed instability into compensatory counterrotation of the eyes. Sensory feedback tunes this vestibulo-ocular reflex throughout life. We studied the functional development of vestibulo-ocular reflex circuit components in the larval zebrafish, with and without sensation. Blind fish stabilize gaze normally, and neural responses to body tilts mature before behavior. In contrast, synapses between motor neurons and the eye muscles mature with a time course similar to behavioral maturation. Larvae without vestibular sensory experience, but with mature neuromuscular junctions, had a strong vestibulo-ocular reflex. Development of the neuromuscular junction, and not sensory experience, therefore determines the rate of maturation of an ancient behavior.**

Early sensory deprivation profoundly disrupts neural function and associated behaviors in visual (*1*), auditory (*2*), and vestibular (*3*, *4*) systems, and sensory feedback is responsible for tuning mature vestibular behaviors (*5*). Sensory experience might therefore set the pace at which neural circuits and the behaviors they subserve emerge; alternatively, such functional maturation might reflect development of the underlying circuit components. Although each component of a neural circuit plays a necessary role in generating behavior, a component's necessity does not confer rate-limiting status. Instead—by analogy to chemistry (*6*)—a rate-limiting component would be one whose development is the most protracted, determining the rate at which behavior can reach maturation. Despite recent technological improvements in connectomics for circuit identification (*7*, *8*), the complexity and in utero development (*9*) of mammalian circuits complicates measuring rates of development and their link to specific behaviors.

We studied an archetypal sensorimotor circuit that stabilizes gaze in a simple model vertebrate, the larval zebrafish. The relative simplicity and high conservation of this circuit across vertebrates (*10*) have made larval zebrafish a powerful model to uncover neural mechanisms of sensorimotor behavior (*11*). The vestibulo-ocular reflex circuit consists of sensory afferents, central interneurons, and motor neurons that together transform head and

---

[1] Department of Otolaryngology, Department of Neuroscience and Physiology, and the Neuroscience Institute, NYU Grossman School of Medicine, New York, NY, USA. [2] Center for Neural Science, New York University, New York, NY, USA.
***Corresponding author. Email: paigebelleleary@gmail.com (P.L.); schoppik@gmail.com (D.S.)**

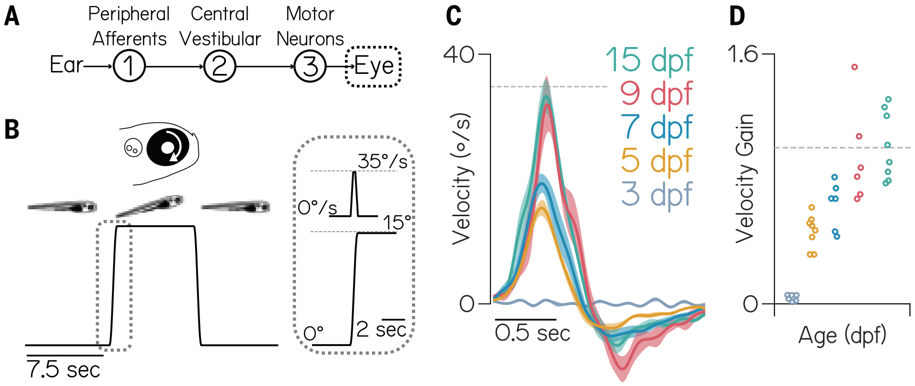


**Fig. 1. Gaze stabilization matures by 9 days postfertilization.** (**A**) Diagram of the feed-forward gaze stabilization circuit. Tilts are sensed by the utricle in the inner ear; information then flows through the eighth cranial nerve peripheral afferents (1) to central vestibular neurons in the brainstem (2) to motor neurons in cranial nuclei nIII and nIV (3) that counterrotate the eye. (**B**) Schematic of body tilt stimulus. Gray dotted box shows the trapezoidal velocity profile and angle of the tilt. (**C**) Average eye velocity ± SEM at 3, 5, 7, 9, and 15 dpf ($N$ = 6, 8, 6, 6, and 8 fish). Gray line at the peak tilt velocity (35°/s). (**D**) Gain (peak eye velocity divided by peak table velocity) for each fish (peak table velocity = 35°/s); color indicates age as in (C). Three versus 5 dpf gains, $P_{\text{post hoc}} = 0.0021$; 3 versus 7 dpf gains, $P_{\text{post hoc}} = 6.51 \times 10^{-5}$; 3 versus 9 dpf gains, $P_{\text{post hoc}} = 4.57 \times 10^{-8}$; 9 versus 15 dpf gains, $P_{\text{post hoc}} = 0.95$. Horizontal line at gain = 1.

body tilts into counterrotatory eye movements (*12*) (Fig. 1A). When mature, this feed-forward circuit generates eye movements matching the head and body velocity, minimizing retinal slip and stabilizing gaze. In vertebrates, both gaze stabilization and vestibulo-ocular reflex circuit components mature during early development (*13*). In this study, we used the transparent, externally developing larval zebrafish to determine the role of sensation in vestibulo-ocular reflex circuit development.

## Early maturation of gaze stabilization behavior

To determine when behavior matures, we measured the eyes' response to body tilts across early development (Fig. 1B). We previously used this approach to define the frequency response of the vestibulo-ocular reflex (*14*). Briefly, fish were immobilized with their eye freed and tilted in the pitch axis (nose-up or nose-down) on a rotating platform in complete darkness. Tilt magnitude (15°) and peak velocity (35°/s) were selected to match the statistics of pitch tilts observed in freely

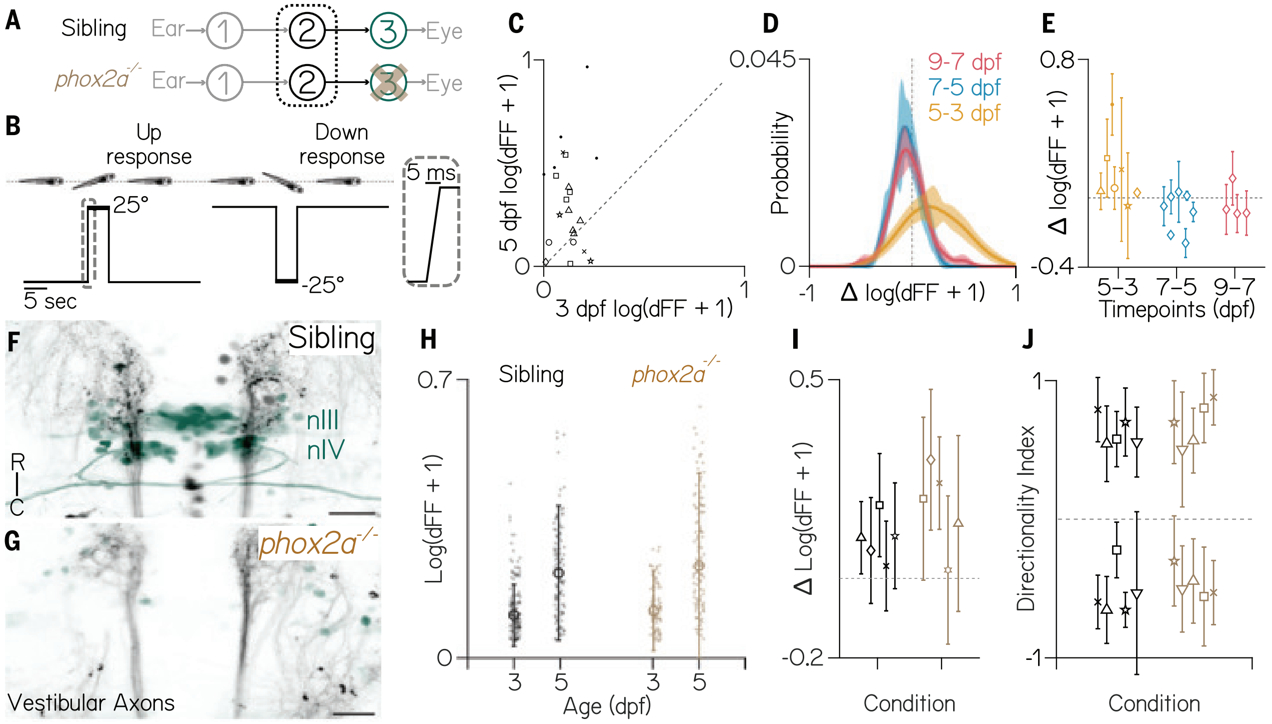


**Fig. 2. Central vestibular neuron responses plateau between 3 and 5 dpf, with or without motor neurons.** (**A**) Diagram of the gaze stabilization circuit, focused on vestibular neurons in fish with (green) and without (brown, *phox2a*) motor neurons in cranial nuclei nIII and nIV. (**B**) A pitch-tilt stimulus trial with a 15-s baseline, a rapid step (inset, 25° step), a 5-s eccentric hold, and a rapid return for imaging. (**C**) Comparison of neural responses (the change in fluorescence divided by basal fluorescence, or dFF) to 25° nose-up steps at 3 and 5 dpf ($P_{t\text{ test}} = 1.2 \times 10^{-3}$). Dashed line at 0. (**D**) Distribution ± bootstrapped SD of pairwise differences of individual neurons between days (5 and 3, 7 and 5, and 9 and 7 dpf) in response to 25° steps. (**E**) Data from (D), plotted as the median pairwise difference ± interquartile ratio (IQR) for neurons from each individual fish ($N$ = 7, 7, and 4). Five to 7 versus 7 to 9 dpf ($P_{t\text{ test}}$ = 0.067 and 0.22, respectively). Dashed line at 0. (**F** and **G**) Axons of central vestibular nuclei (black) to motor neurons in nIII and nIV (green) in a 3 dpf sibling and *phox2a* mutant. Scale bars: 25 μm. R, rostral; C, caudal. (**H**) Individual neuron responses (dots) to 19° pitch tilts in their preferred direction at 3 and 5 dpf in *phox2a*$^{-/-}$ mutants and controls ($P_{\text{ANOVA}}$ for response change over time $< 10 \times 10^{-16}$; $P_{\text{ANOVA}}$ for age and genotype interactions $< 10 \times 10^{-16}$). Median ± IQR overlaid. (**I**) Data from (H), plotted as the median pairwise difference ± IQR for paired neurons from each individual fish ($n$ = 113 neurons/$N$ = 5 siblings, $n$ = 116/$N$ = 5 *phox2a* mutants; $P_{\text{ANOVA}} = 5.03 \times 10^{-9}$ for age effect, $P_{\text{ANOVA}}$ = 0.068 for interactions of age and genotype). Dashed line at 0. (**J**) Directionality indices of vestibular neurons in *phox2a*$^{-/-}$ mutants ($N$ = 5 fish, $n$ = 102 neurons) and controls ($N$ = 5, $n$ = 112), plotted as the median ± IQR for each fish ($P_{\text{K-S test}}$ = 0.66).

swimming larvae (*15*). At and after 5 days postfertilization (dpf), fish rotated their eyes down in response to nose-up pitch tilt stimuli (Fig. 1C). A mature vestibulo-ocular reflex will counterrotate the eyes at the same velocity as the body and head, eliminating residual image motion (retinal slip) and stabilizing gaze. We therefore parameterized behavioral performance as the ratio of the peak eye velocity to the peak platform velocity (gain). A gain of 1 is fully compensatory and mature, 0 indicates that the eye does not move, and gains between 0 and 1 reflect immature and inadequate gaze stabilization. Gain improved with age from 3 to 9 dpf ($P_{\mathrm{ANOVA}} = 1.01 \times 10^{-9}$; ANOVA, analysis of variance; Fig. 1D), reached gains of ~1, and did not change between 9 and 15 dpf ($P_{\mathrm{post\ hoc}} = 0.95$). These data indicate that behavior improves gradually over the first week of life and performance plateaus at ~9 dpf.

### Gaze stabilization develops normally in congenitally blind fish

As the vestibulo-ocular reflex minimizes retinal slip, we first sought to determine how the absence of visual feedback during early life would affect the development of gaze stabilization behavior. Visual feedback is particularly important because fish ocular muscles do not contain proprioceptors (Golgi tendon organs and muscle spindles). We raised congenitally blind fish to ~9 dpf (see methods in the supplementary materials for details) and measured eye rotations in response to body tilts, as in Fig. 1 (fig. S1). Blind fish and sighted siblings showed similar gain ($P_{\mathrm{ANOVA}} = 0.934$). Despite its importance for calibrating gain in mature animals, visual information is therefore dispensable for the development of the vestibulo-ocular reflex.

### Vestibular neuron responses plateau well before behavioral performance peaks, with or without motor neuron–derived feedback

Given that visual feedback is dispensable, either the slowest component of the circuit to develop

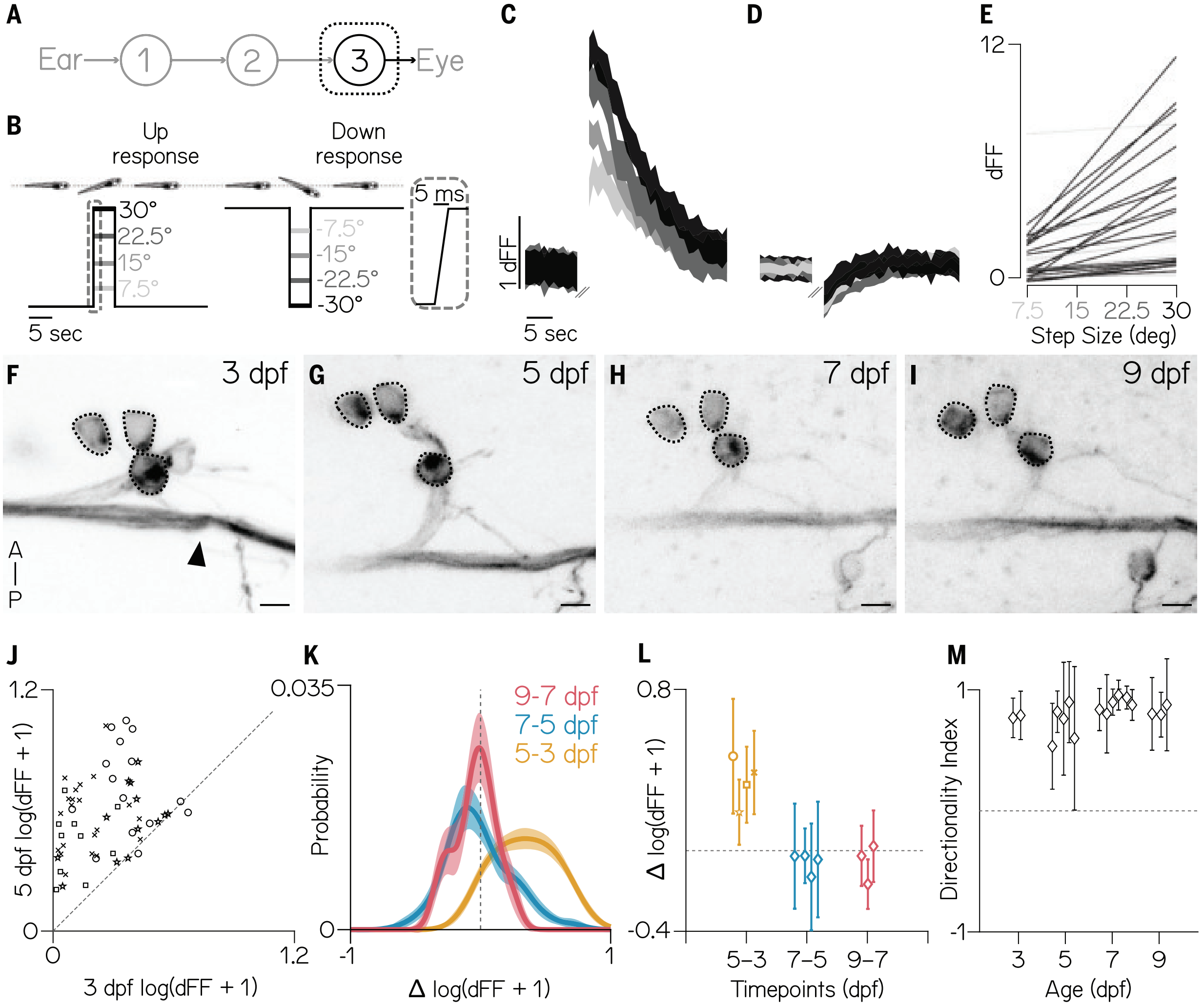


**Fig. 3. Superior oblique motor neuron responses plateau between 3 and 5 dpf.** (**A**) Diagram of the gaze stabilization circuit, focused on motor neurons. (**B**) A pitch-tilt stimulus trial with a 15-s baseline, a rapid step (inset, 30° step), a 5-s eccentric hold, and a rapid return for imaging. (**C** and **D**) Mean ± SD of responses (dFF) to nose-up (C) and nose-down (D) pitch steps from a single superior oblique extraocular motor neuron at 5 dpf (three to five trials per step size). Pitch amplitude (grayscale) as in (B). (**E**) Best-fit lines of responses (dFF) from superior oblique motor neurons to nose-up pitch tilts. Black lines have slopes > 0 ($n = 22/32$, $N = 3$ fish). (**F** to **I**) Three superior oblique extraocular motor neurons (dotted circles) and the trochlear nerve (black arrowhead) tracked from 3 to 9 dpf. Scale bars: 5 μm. A, anterior; P, posterior. (**J**) Comparison of neural responses (dFF) to 25° nose-up steps at 3 and 5 dpf ($P_{t\ \mathrm{test}} = 1.31 \times 10^{-16}$). Dashed line at 0. (**K**) Distribution ± bootstrapped SD of pairwise differences of individual neurons between days (5 and 3, 7 and 5, 9 and 7 dpf) in response to 25° nose-up steps. (**L**) Data from (K), plotted as the median pairwise difference for neurons from each individual fish ($N = 4$, 4, and 3). Five to 7 versus 7 to 9 dpf ($P_{t\ \mathrm{test}} = 0.69$ and 0.23, respectively). Dashed line at 0. (**M**) Median ± IQR directionality index to ±25° steps across ages ($P_{\mathrm{K\text{-}W\ test}} = 0.09$). +1 is selective for nose-up, −1 is selective for nose-down. Dashed line at 0.

and/or nonvisual feedback could constrain behavioral maturation. Vestibular interneurons make up the central node of a feed-forward circuit, where their activity encodes body and head tilt magnitude and direction (Fig. 2A). To measure the development of vestibular responses to body tilts, we recorded neural activity at eccentric body postures (Fig. 2B).

We used Tilt In Place Microscopy (TIPM) (*16*) to avoid rotating the microscope. TIPM returns fish from an eccentric orientation to the imaging plane faster (~5 ms) than the time constant of the calcium indicator (GCaMP6s) (*17*). The activity observed reflects the decay of the neuron's eccentric response upon return to horizontal. Central neurons in the tangential vestibular nucleus that project to extraocular motor nuclei nIII and nIV are indispensable for the vestibulo-ocular reflex (*14*, *18*). They respond exclusively to either nose-up or nose-down tilts (*19*), and, when activated optogenetically, they induce torsional eye rotations (*18*). We performed longitudinal TIPM (±25°) in these neurons to determine when their responses plateau. We found that vestibular neuron response strength increases drastically between 3 and 5 ($P_{t\ \mathrm{test}} = 1.2 \times 10^{-3}$, $n = 22$ neurons, $N = 7$ fish) but not between 5 and 7 or 7 and 9 dpf ($P_{t\ \mathrm{test}} = 0.067$, $n = 21$ neurons, $N = 7$ fish; Fig. 2, C to E). Although both behavior and vestibular neuron responses improved between 4 and 5 dpf, vestibular neuron response strength reached a plateau well before behavioral performance is mature. The slowest component to mature must therefore be downstream of central vestibular interneurons.

To determine whether motor-derived feedback influences the maturation of vestibular neuron responses, we adopted a loss-of-function approach. We performed TIPM to measure vestibular neuron responses on a *phox2a* mutant background (*20*) that fails to develop nIII and nIV extraocular motor neurons (Fig. 2, A, F, and G). After ±19° tilts, responses of individual vestibular neurons (Fig. 2H) increased between 3 and 5 dpf for both sibling fish and mutants ($P_{\mathrm{ANOVA}} = 1.95 \times 10^{-26}$, age × genotype $P_{\mathrm{ANOVA}} = 0.053$; 3 dpf: $n = 137$ neurons over $N = 5$ siblings, $n = 88$ neurons/$N = 5$ mutants; 5 dpf: $n = 113$ neurons/$N = 5$ siblings, $n = 116$ neurons/$N = 5$ mutants). Across fish (Fig. 2I), the average response of paired (tracked) neurons increased between 3 and 5 dpf ($P_{\mathrm{ANOVA}} < 10 \times 10^{-16}$) with no effect of genotype ($P_{\mathrm{ANOVA}} = 0.069$). Finally, to evaluate the development of peripheral vestibular inputs, we examined directional selectivity at 5 dpf. Consistent with a previous report (*19*), the ratio of nose-up and nose-down sensitive (directionality index values < |0.1|) neurons was nearly even, with no difference between genotype ($P_{\mathrm{K\text{-}S\ test}} = 0.66$; 51.9 ± 6% nose-up, 48 ± 6% nose-down, $n = 118$ neurons over $N = 5$ siblings; 48.1 ± 13% nose-up stimuli, 51.8 ± 13% nose-down, $n = 117$/$N = 5$ *phox2a* mutants) (Fig. 2J). These results establish that, even without motor neurons—and, by extension, motor-derived feedback—vestibular neuron responses develop early relative to behavior.

## Extraocular motor neuron responses develop between 3 and 5 dpf

As the last central node in a feed-forward reflex, extraocular motor neuron activity reflects all sensory inputs from the entire circuit (Fig. 3A). Consequentially, if the development of motor neuron responses reaches a plateau earlier than behavioral improvement, the slowest

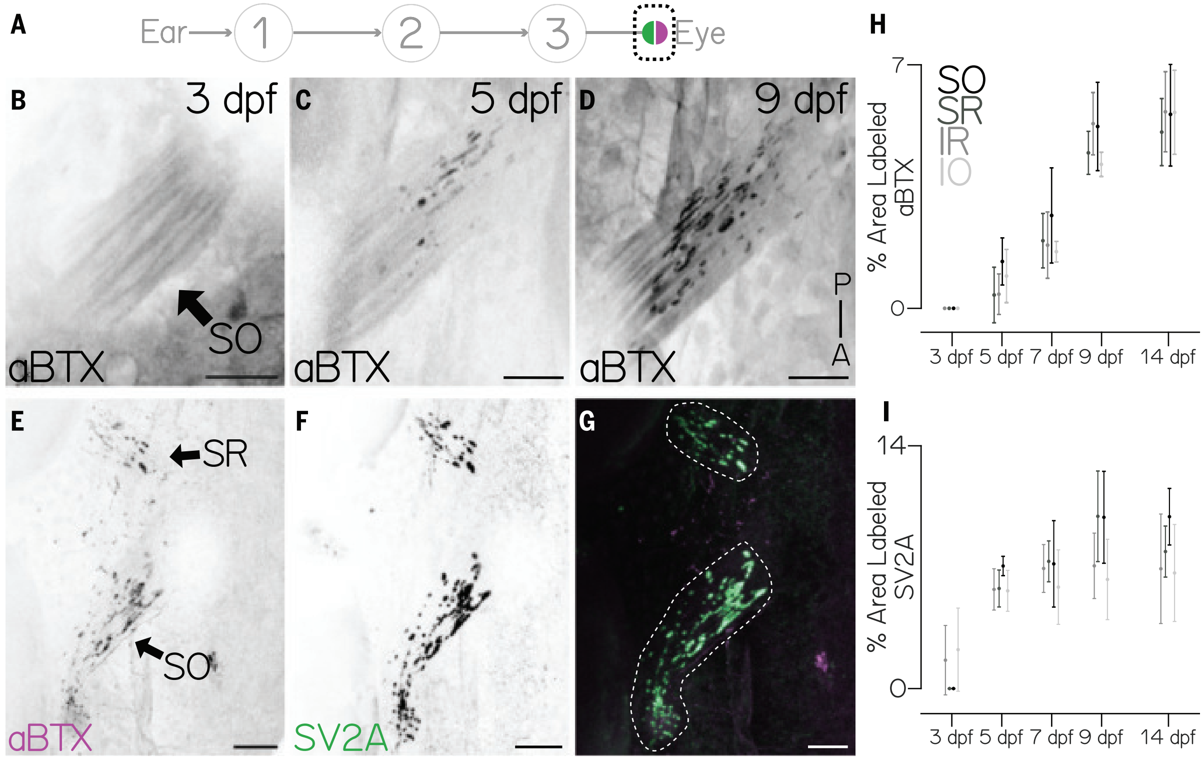


**Fig. 4. The time course of postsynaptic neuromuscular junction development matches behavioral maturation.** (**A**) Diagram of the gaze stabilization circuit, focusing on presynaptic (green) and postsynaptic (magenta) components of the neuromuscular junction. (**B** to **D**) Dorsal projection of the superior oblique muscle (black arrow) at 3 (B), 5 (C), and 9 (D) dpf labeled with fluorescent α-BTX. Scale bars: 20 μm. (**E** to **G**) Comparison of superior oblique (SO, arrow) and superior rectus (SR, arrow) presynaptic (SV2A, green) and postsynaptic (α-BTX, magenta) label at 8 dpf. Dashed line in (G) denotes muscle bound used for quantification. SO muscle: 6.59% α-BTX and 15.03% SV2A. SR muscle: 4.98% α- BTX and 9.33% SV2A. Scale bars: 20 μm. (**H**) Postsynaptic α-BTX label in SO, SR, inferior rectus (IR), and inferior oblique (IO) muscles. Dots are the mean ± SD area labeled. Three versus 9 dpf, $P_{\mathrm{ANOVA}} < 10 \times 10^{-16}$; 9 versus 14 dpf, $P_{\mathrm{post\ hoc}} = 0.185$. (**I**) Presynaptic SV2A staining in SO, SR, IR, and IO. Dots are the mean ± SD area labeled. Three versus 5 dpf, $P_{\mathrm{post\ hoc}} = 5.5 \times 10^{-4}$; 5 versus 7 dpf, $P_{\mathrm{post\ hoc}} = 0.889$; 7 versus 9 dpf, $P_{\mathrm{post\ hoc}} = 0.462$; 9 versus 14 dpf, $P_{\mathrm{post\ hoc}} = 0.983$.

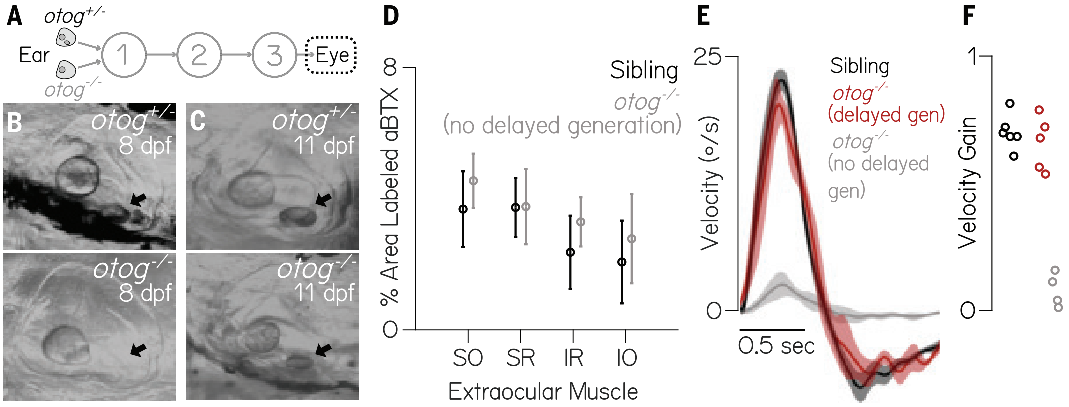


**Fig. 5. Gaze stabilization behavior emerges rapidly after delayed emergence of tilt sensation.** (**A**) Diagram of the gaze stabilization circuit illustrating *otogelin*$^{-/-}$ phenotype with focus on eye movements. (**B** and **C**) (Top) *otogelin*$^{+/-}$, utricular otolith visible in ear at 8 (left) and 11 dpf (right, black arrows). (Bottom) *otogelin*$^{-/-}$ fish without a utricular otolith at 8 dpf and with at 11 dpf (black arrows). (**D**) Postsynaptic α-BTX label from sibling and *otogelin*$^{-/-}$ fish in SO ($N$ = 6 and 6), SR ($N$ = 5 and 6), IR ($N$ = 5 and 5), and IO ($N$ = 4 and 6). (**E**) Average eye velocity traces ± SEM to a nose-up step as in Fig. 1B. (**F**) Vestibulo-ocular reflex gain (peak eye velocity divided by peak table velocity) for each fish (peak table velocity = 35°/s) ($N$ = 6, 5, and 4) in (E). Gains in *otogelin*$^{-/-}$ fish with delayed utricular generation versus controls, $P_{\text{post hoc}} = 0.66$; gains in *otogelin*$^{-/-}$ fish with no utricular generation versus controls, $P_{\text{post hoc}} = 1.84 \times 10^{-7}$; gains in *otogelin*$^{-/-}$ fish with no utricular generation versus *otogelin*$^{-/-}$ fish with delayed utricular generation, $P_{\text{post hoc}} = 6.20 \times 10^{-7}$.

developmental process must be in the motor periphery. We repeated longitudinal TIPM (±25°) to determine when motor neuron responses plateau. However, as the calcium responses of motor neurons have not been evaluated in response to body tilts, we first characterized the tuning and sensitivity properties of these neurons to a gradient of stimulus steps (Fig. 3B). To match the pulling direction of the muscle, superior oblique motor neurons should respond predominantly to nose-up tilts. As expected, superior oblique neurons at 5 dpf responded unidirectionally to nose-up body tilts (Fig. 3C) but not to nose-down body tilts (Fig. 3D). Motor neuron activity at 5 dpf varied as a function of nose-up tilt eccentricity ($n$ = 22/32 neurons with slope > 0, $P_{\text{post hoc}} = 0.002$; Fig. 3E).

To measure how tilt responses developed, we recorded calcium signals from a transgenic line with sparse expression in superior oblique motor neurons (fig. S2). Motor neuron somata were stably positioned within nIV (Fig. 3, F to I), allowing reliable identification of the same neurons over 2-day increments: 3 to 5 dpf, 5 to 7 dpf, and 7 to 9 dpf. Almost every superior oblique neuron had a stronger response to 25° nose-up body tilts at 5 dpf than at 3 dpf ($P_{t\text{ test}} = 1.31 \times 10^{-16}$, $n$ = 59 pairs over $N$ = 4 fish; Fig. 3J). Whereas the distributions of differences varied between 5 to 7 and 7 to 9 dpf ($P_{\text{K-S test}} = 5.7 \times 10^{-15}$; Fig. 3K), when evaluated across fish, the average response did not change between 5 to 7 dpf ($P_{t\text{ test}} = 0.69$, $n$ = 53 cells/$N$ = 4 fish) or 7 to 9 dpf ($P_{t\text{ test}} = 0.23$, $n$ = 23/$N$ = 3; Fig. 3L). We gathered another dataset with 19° nose-up and nose-down steps to evaluate the development of directional selectivity (see methods for details). We observed no changes across time ($P_{\text{K-W test}} = 0.09$, 3 dpf: $n$ = 16 neurons over $N$ = 2 fish; 5 dpf: $n$ = 32/$N$ = 5; 7 dpf: $n$ = 60/$N$ = 7; 9 dpf: $n$ = 30/$N$ = 3; Fig. 3M). Brain activity related to the vestibulo-ocular reflex (*21*) must converge on extraocular motor neurons. Because extraocular motor neuron response strength and directional selectivity appear to plateau by 5 dpf—well before behavior—we conclude that the slowest component of the circuit to develop must lie downstream of motor neurons.

## The developmental time course of the postsynaptic neuromuscular junction matches the maturation of behavior

We next focused on the extraocular neuromuscular junction (Fig. 4A). We labeled postsynaptic acetylcholine receptors with fluorescently conjugated alpha-bungarotoxin (α-BTX; Fig. 4, B to D) and presynaptic vesicles with an SV2A antibody (*22*) (Fig. 4, E to G). We targeted all four eye muscles used for torsional gaze stabilization: superior oblique (SO), superior rectus (SR), inferior oblique (IO), and inferior rectus (IR) in 3 to 9 and 14 dpf fish. Over time, the fraction of each muscle labeled with α-BTX increased with a time course comparable to behavioral improvement ($P_{\text{ANOVA}} < 10 \times 10^{-16}$) (SO, $n$ = 106 muscles over $N$ = 61 fish; SR, $n$ = 52/$N$ = 32; IR, $n$ = 43/$N$ = 28; IO, $n$ = 38/$N$ = 26; Fig. 4H), with no difference in labeling observed between 9 and 14 dpf ($P_{\text{post hoc}} = 0.185$). In contrast, SV2A labeling appeared earlier and developed more rapidly. SV2A labeling significantly increased between 3 and 5 dpf ($P_{\text{post hoc}} = 5.5 \times 10^{-4}$) but did not change afterward ($P_{\text{post hoc}} > 0.13$ for 5 to 7, 5 to 9, 5 to 14, 7 to 9, 7 to 14, and 9 to 14; SO, $n$ = 34 cells/$N$ = 17 fish; SR, $n$ = 33/$N$ = 17; IR, $n$ = 29/$N$ = 17; IO, $n$ = 28/$N$ = 15; Fig. 4I).

Larval zebrafish eye muscles assume their adult configuration by 3 dpf and develop both thick and thin myofibrils by 4 dpf (*23*), allowing the eye to assume static orientations up to ±30° (*14*) and a mature optokinetic response (<0.5 Hz) by 6 dpf (*24*). To evaluate whether eye velocity was constrained by the properties of the muscle, we measured eye movements at 5 dpf after unnaturally rapid tilts (90°/s, 600°/s$^2$ instead of 35°/s, 150°/s$^2$; fig. S3A). We found that although gaze stabilization behavior was far from mature, and the α-BTX signal had just begun to emerge, the eye could rotate faster ($P_{t\text{ test}} = 0.016$; fig. S3, B and C). Together, these findings implicate postsynaptic development at the neuromuscular junction as the slowest step in circuit maturation.

## Vestibulo-ocular reflex behavior is mature after restoration of transient sensory deprivation

If sensory experience sets the rate of behavioral development, then transient loss of vestibular sensation should delay normal improvements to gaze stabilization. Upon restoration, vestibulo-ocular reflex performance should steadily improve. Alternatively, if behavioral performance reflects the capacity of the circuit component that matures most slowly, then upon restoration of vestibular sensation, behavior would be immediately comparable between experienced and transiently deprived larvae.

We investigated the emergence of the vestibulo-ocular reflex in a mutant line (*otogelin*) that, in some cases (*25*), is transiently insensitive to body tilts. Under normal conditions (*19*), larval zebrafish rely exclusively on the utricle to sense body tilts (*26*), which uses an inertial difference between an otolith and hair cells embedded in a gelatinous macula to transduce linear acceleration. Initial calcification of the utricular otolith normally occurs between 18 and 24 hours postfertilization (*27*). Most *otogelin* mutants do not generate a utricular otolith and as a result, in the dark, cannot sense body tilts (*28*). A small fraction of *otogelin* mutants show delayed generation of the utricular otolith, which appears over a 24-hour period at around 2 weeks of age (Fig. 5, B and C), at which point postural behaviors resume (*25*).

We hypothesized that because the neuromuscular junction would have time to mature, *otogelin* mutants should show comparable behavior to siblings as soon as the utricle becomes functional.

We first confirmed that the neuromuscular junction developed normally in *otogelin* mutants; we observed strong α-BTX labeling at 8 dpf in both mutants and sibling controls in all four muscles (Fig. 5D). Next, we raised and screened >2000 *otogelin* mutants daily to identify five fish with nascent bilateral utricular otoliths between 11 and 16 dpf (see methods for details). Eye movements in response to body tilts were measured (Fig. 1B) on the day after identification, when the otolith had reached normal size. Performance was statistically indistinguishable between controls ($N = 6$) and mutants ($N = 5$) with newly generated otoliths ($P_{\text{post hoc}} = 0.66$; Fig. 5, E and F). In contrast, 11 to 16 dpf fish that never developed otoliths were profoundly impaired ($P_{\text{post hoc}} = 1.84 \times 10^{-7}$ sibling versus no otolith, $6.20 \times 10^{-7}$ delayed otolith versus no otolith). We infer from the rapid emergence of functional gaze stabilization in *otogelin* mutants that initial vestibular circuit assembly can proceed without sensory input, consistent with other reports (*25*, *29*, *30*). Our data show that vestibular experience is dispensable for maturation of the vestibulo-ocular reflex.

## Discussion

We find that both the vestibulo-ocular reflex circuit and behavior can mature without any vestibular or visual input. This finding is particularly notable given that gaze stabilization is plastic later in life: Visual feedback from eye movements is used to adjust the sensitivity of central vestibular neurons to modulate gain (*5*). Even when there are no motor neurons to move the eyes, we see that central vestibular neuron responses to tilts are unchanged, underscoring the early dispensability of sensory feedback. In the light, the vestibulo-ocular and optokinetic reflexes work together to stabilize gaze; feedback from optokinetic eye movements might suffice to shape the vestibulo-ocular reflex circuit we studied here. However, such modulation would be inconsistent with our finding that the vestibulo-ocular reflex matures normally in congenitally blind fish. Studies of locomotor development over the past century underscore the importance of feedback as animals learn to move properly in their environment (*31*, *32*). Undoubtedly, given the considerable morphological and neurological changes that happen between larval and adult stages, feedback will be similarly important to maintain an excellent vestibulo-ocular reflex. We show that such plasticity acts on a mature scaffold that can evoke excellent behavior as soon as the neuromuscular junction is sufficiently developed.

By definition, evolutionary adaptations that support earlier or later development of sensorimotor reflex behaviors must manifest with changes at the circuit component that is slowest to develop—that is, a rate-limiting step. Does the rate of development of the neuromuscular junction limit maturation of other behaviors? Early neuromuscular junction development permits developing rodents to make spontaneous limb contractions, or "twitches," that can shape sensory cortical (*33*) and cerebellar development (*34*). Similar twitches (*35*) precede development of the hindbrain neurons responsible for evoked escape swimming in larval zebrafish (*36*), the earliest sensorimotor reflex behavior; by inference, the rate-limiting process for escapes is upstream of the motor periphery. Differences in the location of the rate-limiting process may reflect selective pressures that differ between precocial (mobile at birth) and altricial (immobile) animals. Empirically, when the pressure is on to stand (or run or swim) or be eaten, the motor periphery develops much more rapidly to ensure that "the neuromuscular system [is] ready for use when the brain needs it" (*37*). We suggest that defining the location of rate-limiting processes will provide critical insights into how brain circuits develop.

Recent advances in transcriptomics and connectomics (*7*, *8*) revealed the elements of and interactions between vertebrate neural circuit nodes. Similar advances transformed longitudinal (*38*) and circuit-level (*39*) measurements linking neuronal activity and behavior (*40*). In contrast, conceptual frameworks (*41*) lag behind, relying primarily on classic loss-of-function approaches to establish necessity (*42*). Here, we propose a candidate rate-limiting process for the maturation of a vertebrate sensorimotor behavior. Our approach moves beyond "necessity" (*43*), linking neural circuit development to behavioral performance.

### REFERENCES AND NOTES

1. D. H. Hubel, T. N. Wiesel, *Naunyn Schmiedebergs Arch. Pharmakol.* **248**, 492–497 (1964).
2. E. F. Chang, M. M. Merzenich, *Science* **300**, 498–502 (2003).
3. M. Jamon, *Front. Integr. Neurosci.* **8**, 11 (2014).
4. C. T. Goode, D. L. Maney, E. W. Rubel, A. F. Fuchs, *J. Neurophysiol.* **85**, 1119–1128 (2001).
5. S. du Lac, J. L. Raymond, T. J. Sejnowski, S. G. Lisberger, *Annu. Rev. Neurosci.* **18**, 409–441 (1995).
6. F. F. Blackman, *Ann. Bot.* **os-19**, 281–296 (1905).
7. J. Zhou *et al.*, *Nature* **624**, 355–365 (2023).
8. L. Yuan, X. Chen, H. Zhan, G. L. Henry, A. M. Zador, *Nat. Commun.* **15**, 8371 (2024).
9. M. E. Thomason, *Biol. Psychiatry* **88**, 40–50 (2020).
10. H. Straka, R. Baker, *Front. Neural Circuits* **7**, 182 (2013).
11. J. M. Goldberg *et al.*, *The Vestibular System: A Sixth Sense* (Oxford Univ. Press, 2012).
12. J. Szentágothai, *J. Neurophysiol.* **13**, 395–407 (1950).
13. M. Beraneck, F. M. Lambert, S. G. Sadeghi, in *Development of Auditory and Vestibular Systems*, R. Romand, I. Varela-Nieto, Eds. (Elsevier, ed. 4, 2014), pp. 449–487.
14. I. H. Bianco *et al.*, *Curr. Biol.* **22**, 1285–1295 (2012).
15. Y. Zhu *et al.*, *Cell Rep.* **42**, 112573 (2023).
16. K. R. Hamling, Y. Zhu, F. Auer, D. Schoppik, *J. Neurosci.* **43**, 936–948 (2023).
17. T.-W. Chen *et al.*, *Nature* **499**, 295–300 (2013).
18. D. Schoppik *et al.*, *J. Neurosci.* **37**, 11353–11365 (2017).
19. D. Goldblatt *et al.*, *Curr. Biol.* **33**, 1265–1281.e7 (2023).
20. D. Goldblatt *et al.*, *eLife* **13**, RP96893 (2024).
21. G. Migault, N. Beiza-Canelo, S. Chatterjee, G. Debrégeas, V. Bormuth, bioRxiv 2024.03.22.586054 [Preprint] (2024); https://doi.org/10.1101/2024.03.22.586054.
22. K. Buckley, R. B. Kelly, *J. Cell Biol.* **100**, 1284–1294 (1985).
23. S. S. Easter Jr., G. N. Nicola, *Dev. Psychobiol.* **31**, 267–276 (1997).
24. J. C. Beck, E. Gilland, D. W. Tank, R. Baker, *J. Neurophysiol.* **92**, 3546–3561 (2004).
25. R. Roberts, J. Elsner, M. W. Bagnall, *J. Assoc. Res. Otolaryngol.* **18**, 415–425 (2017).
26. B. B. Riley, S. J. Moorman, *J. Neurobiol.* **43**, 329–337 (2000).
27. B. B. Riley, C. Zhu, C. Janetopoulos, K. J. Aufderheide, *Dev. Biol.* **191**, 191–201 (1997).
28. T. T. Whitfield *et al.*, *Development* **123**, 241–254 (1996).
29. F. Ulrich, C. Grove, J. Torres-Vázquez, R. Baker, *Curr. Neurobiol.* **7**, 62–73 (2016).
30. D. L. Barabási, G. F. P. Schuhknecht, F. Engert, *Nat. Commun.* **15**, 364 (2024).
31. K. E. Adolph, B. I. Bertenthal, S. M. Boker, E. C. Goldfield, E. J. Gibson, *Monogr. Soc. Res. Child Dev.* **62**, i–162 (1997).
32. E. Thelen, *Int. J. Behav. Dev.* **24**, 385–397 (2000).
33. R. Khazipov *et al.*, *Nature* **432**, 758–761 (2004).
34. J. C. Dooley, G. Sokoloff, M. S. Blumberg, *Curr. Biol.* **31**, 5501–5511.e5 (2021).
35. L. Saint-Amant, P. Drapeau, *J. Neurobiol.* **37**, 622–632 (1998).
36. R. C. Eaton, R. D. Farley, *Copeia* **1973**, 673–682 (1973).
37. D. Kernell, *Neurosci. Biobehav. Rev.* **22**, 479–484 (1998).
38. N. A. Steinmetz *et al.*, *Science* **372**, eabf4588 (2021).
39. T. Nöbauer, Y. Zhang, H. Kim, A. Vaziri, *Nat. Methods* **20**, 600–609 (2023).
40. S. Schneider, J. H. Lee, M. W. Mathis, *Nature* **617**, 360–368 (2023).
41. D. L. Barabási *et al.*, *J. Neurosci.* **43**, 5989–5995 (2023).
42. Y. Lazebnik, *Cancer Cell* **2**, 179–182 (2002).
43. S. Brenner, *Curr. Biol.* **5**, 332 (1995).
44. D. Schoppik, Data associated with "Sensation is Dispensable for the Maturation of the Vestibulo-ocular Reflex," Open Science Foundation (2024); https://doi.org/10.17605/OSF.IO/7Z5UP.

### ACKNOWLEDGMENTS

We thank C. Moens and M. Bagnall for all the fish; H. Gelnaw for assistance with husbandry; and K. Nagel, J. Dasen, N. Ringstad, L. Kiorpes, M. Long, G. Buzsáki, A. Schier, C. Desplan, S. Burden, and members of the Schoppik and Nagel labs for their valuable feedback and discussions. **Funding:** National Institute on Deafness and Communication Disorders of the National Institutes of Health award number R01DC017489 (D.S.); National Institute on Deafness and Communication Disorders of the National Institutes of Health award number F31DC020910 (P.L.); National Institute for Neurological Disorders and Stroke of the National Institutes of Health award number F99NS129179 (D.G.); National Science Foundation under Graduate Research Fellowship number DGE2041775 (P.L.). **Author contributions:** Conceptualization: P.L. and D.S. Data curation: P.L. Formal analysis: P.L. Funding acquisition: P.L., D.G., and D.S. Investigation: P.L., C.B., and C.Q. Methodology: P.L. and D.S. Project administration: D.S. Resources: B.R. and D.G. Software: P.L. and D.S. Supervision: P.L. and D.S. Validation: C.B. Visualization: P.L. and D.S. Writing – original draft: P.L. Writing – review & editing: P.L. and D.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All raw data and analysis code are available at the Open Science Foundation (*44*).

### SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adr9982
Materials and Methods
Figs. S1 to S3
References (*45*–*59*)
MDAR Reproducibility Checklist

Submitted 24 July 2024; accepted 6 November 2024
10.1126/science.adr9982

WILDFIRES

# Canadian forests are more conducive to high-severity fires in recent decades

**Weiwei Wang[1,2], Xianli Wang[1,2]*, Mike D. Flannigan[3], Luc Guindon[4], Tom Swystun[5], Dante Castellanos-Acuna[3,6], Wanli Wu[1,7], Guangyu Wang[1]***

**Canada has experienced more-intense and longer fire seasons with more-frequent uncontrollable wildfires over the past decades. However, the effect of these changes remains unknown. This study identifies driving forces of burn severity and estimates its spatiotemporal variations in Canadian forests. Our results show that fuel aridity was the most influential driver of burn severity, summer months were more prone to severe burning, and the northern areas were most influenced by the changing climate. About 6% (0.54 to 14.64%) of the modeled areas show significant increases in the number of days conducive to high-severity burning during 1981 to 2020, most of which were found during 2001 to 2020 and in the spring and autumn. The extraordinary 2023 fire season demonstrated similar spatial patterns but more-widespread escalations in burn severity.**

As one of the most forested and fire-prevalent countries, Canada is facing longer fire seasons with more-extreme and complex fire behavior driven by changing climate conditions (*1*–*4*). This was exemplified by the record-breaking 2023 fire season, in which ~15 million ha burned—more than seven times the historic average (1986 to 2022) (*5*). However, the assessment of fire effects on forest ecosystems, especially at the national scale, has not been well explored. Burn severity (or fire severity) is widely used as a fire effect metric to measure the changes of aboveground and belowground organic matter caused by fires (*6*, *7*). Burn severity estimates can contribute to assessing postfire vegetation and wildlife responses (*8*, *9*) and to examining fire effects on landscape heterogeneity (*10*) and forest resilience (*11*). Quantifying burn severity potential could play a crucial role in fire risk assessment, management planning, and preparedness strategies to minimize fire-induced costs and losses (*12*, *13*).

From ground surveys, burn severity can be quantified by field-measured continuous indices, such as the composite burn index (CBI) (*6*), burn severity index (BSI) (*14*), and percent overstory mortality (MORT) (*15*). These indices evaluate a single feature of interest or a composite of key factors relevant to the ecological effects of fires to gauge fine-scale burn severity. Categorical burn severity classes can be developed on the basis of these indices, using thresholds derived mostly by relying on ocular judgment and expert knowledge (*6*, *16*). Remote sensing data, on the other hand, provide an alternative estimate of burn severity, especially for large spatial extents and over long time periods. Commonly used remote sensing metrics include the differenced normalized burn ratio (dNBR) (*6*), relativized dNBR (RdNBR) (*16*), and relativized burn ratio (RBR) (*17*). Given their strong associations with field-measured burn severity (*18*–*21*), values of these spectral metrics can be transformed into either categorical classes or continuous indices to approximate field assessments, thereby overcoming the limitations of individual approaches.

A wide range of models have been built to estimate field-measured or remotely sensed burn severity, with random forests (RF) (*22*) being the most used (*23*–*26*). Although adept at modeling nonlinear relationships, machine learning models, such as RF, often struggle with overfitting issues, which hampers their predictive performance. By contrast, simpler linear and generalized linear models, though possibly yielding lower fitting accuracy, may provide greater interpretability and straightforwardness (*15*, *27*). Using these models enables the identification of key drivers of burn severity, with weather or climate, vegetation, and topography typically highlighted as crucial factors, although their relative influences differ among studies (*15*, *24*, *25*).

In Canada, growing efforts have been made to calibrate remotely sensed burn severity measurements (*18*, *19*), generate burn severity products (*20*, *28*), and explore key drivers of burn severity in various regions (*15*, *29*, *30*). However, nationwide burn severity modeling and the identification of its key drivers are still lacking, which renders it impossible to obtain long-term and high–temporal resolution (e.g., daily) burn severity estimates at the national scale. Such estimates are critical to investigate spatial-temporal variations of burn severity potential, which is especially valuable for a country such as Canada, with its extensive and remote forests where wildfires are prevalent but field visits are limited or costly. The estimates may also provide insights into how fire effects vary with changing weather conditions, thereby filling a key knowledge gap about climate change impacts on fire activities (*31*, *32*).

In this work, we built multinomial logistic regression (MLR) models to estimate remotely sensed burn severity for 10 Canadian ecozones (Fig. 1). The dNBR values were categorized into low–, moderate–, and high–burn severity classes on the basis of empirical thresholds suggested by previous studies (*18*–*20*) and used as the dependent variable (table S1). Three groups of predictors were used, including daily weather conditions quantified by the Canadian fire weather index system (CFWIS) indices (*33*, *34*), topographic variation, and fuel characteristics (table S2 and fig. S1). The relative importance of the predictors in the models was derived to identify key drivers of burn severity. MLR models were built using data paired with the daily fire growth data (*35*–*37*) at 30-m resolution from 2001 to 2020 and were then used to estimate daily burn severity by ecozone at 0.25° resolution (~27.8 km) during 1981 to 2020. The estimates presented potential burn severity classes in the event of a fire occurring at each vegetated location. In light of the recent 2023 Canadian extreme fire season, we made additional estimates for 2023 and compared them with the estimates during 1981 to 2020. To investigate spatial and seasonal variations, the number of days conducive to low-, moderate-, or high-severity burning; the monthly potential for high burn severity (the proportion of days per month that were estimated to burn at high severity); and the most severe month (the month with the largest proportion of days that had potential to burn at high severity) in each year (table S3) were summarized for each modeled grid point. To capture the temporal changes of burn severity under recently changing climates, we performed trend analyses for the annual and seasonal summaries of estimated burn severity for the whole 40-year period and two 20-year segments (1981 to 2000 and 2001 to 2020) using the Mann-Kendall test (*38*, *39*) to examine trend significance and the Theil–Sen estimator (*40*, *41*) to estimate the magnitude of change.

## Model performance

The MLR models achieved better accuracy when estimating the high- and low-severity classes in most ecozones, except Hudson plains (HP) (table S4), with a balanced accuracy (a metric of correct predictions equally weighted across classes) of 0.74 (±0.06) and an F1 score (an indicator of precise predictions focusing on a specific class) of 0.66 (±0.08). The performance

---

[1]Department of Forest Resources Management, Faculty of Forestry, The University of British Columbia, Vancouver, BC, Canada. [2]Northern Forestry Centre, Canadian Forest Service, Natural Resources Canada, Edmonton, AB, Canada. [3]Department of Natural Resource Science, Faculty of Science, Thompson Rivers University, Kamloops, BC, Canada. [4]Laurentian Forestry Centre, Canadian Forest Service, Stn. Sainte-Foy, Quebec, QC, Canada. [5]Great Lakes Forestry Centre, Canadian Forest Service, Natural Resources Canada, Sault Ste. Marie, ON, Canada. [6]Department of Renewable Resources, University of Alberta, Edmonton, AB, Canada. [7]Conservation Programs Branch, Parks Canada, Vancouver, BC, Canada.
***Corresponding author. Email: xianli.wang@nrcan-rncan.gc.ca (X.W.); guangyu.wang@ubc.ca (G.W.)**

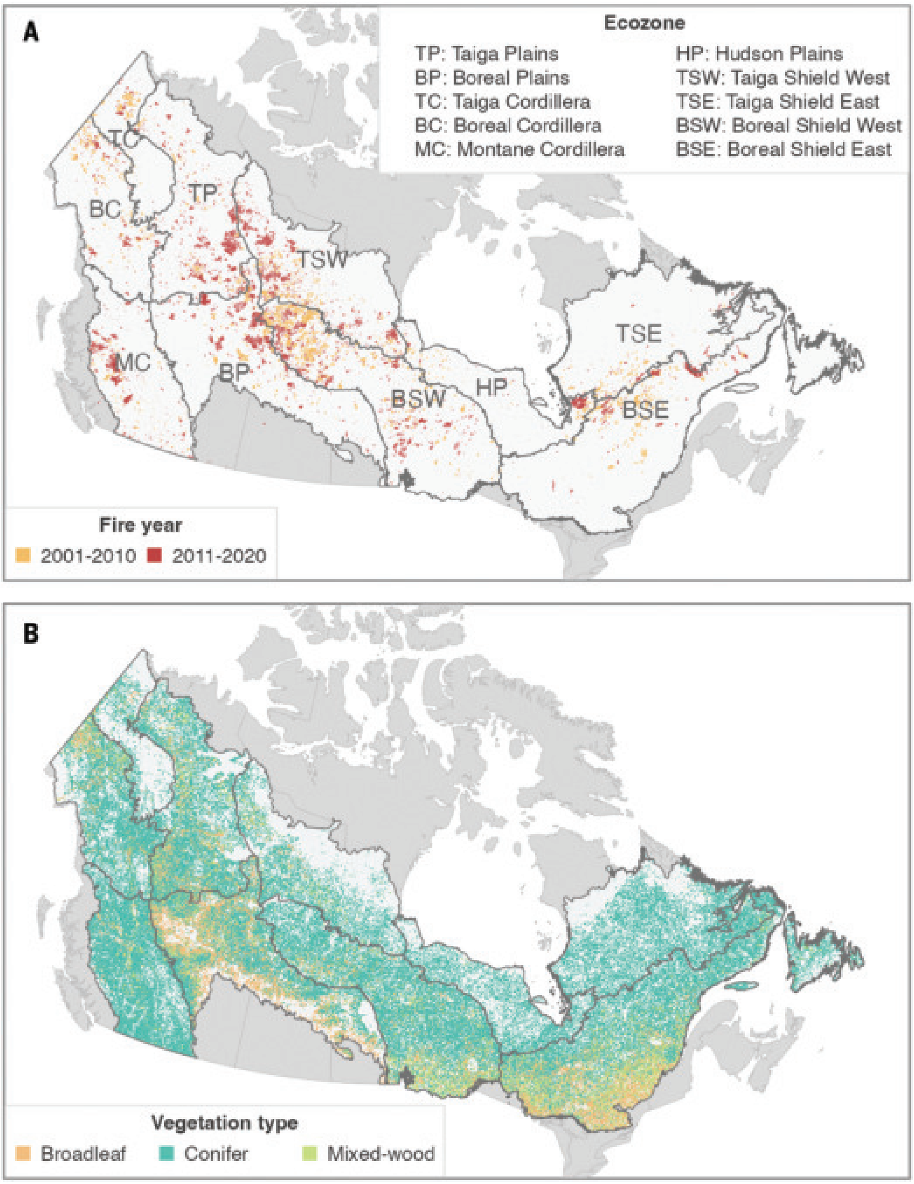


**Fig. 1. Characteristics of the Canadian ecozones.** (**A** and **B**) Fire events during 2001 to 2020 (A) and vegetation type distributions (2020 layer) within the study area (B).

was comparatively lower when estimating the moderate–burn severity class, with a balanced accuracy of 0.67 (±0.06) and an F1 score of 0.56 (±0.09). Among all ecozones, taiga shield west (TSW), taiga cordillera (TC), HP, taiga shield east (TSE), and boreal plains (BP) outperformed the others, with a balanced accuracy of 0.79 (±0.05) and an F1 score of 0.72 (±0.05) for classifying high and low burn severities and a balanced accuracy of 0.72 (±0.03) and an F1 score of 0.63 (±0.05) for classifying moderate burn severity (table S4).

### Fuel aridity determines burn severity in the Canadian forests

Among all individual predictors, two CFWIS indices that measure potential fuel aridity condition, build-up index (BUI) and duff moisture code (DMC), were the most influential for the 10 ecozones (Fig. 2A). Two fuel attribute variables, vegetation type and biomass, also contribute significantly to the models, whereas the most important topographic variables were elevation and slope. Considering the overall contribution of the three predictor groups (weather, topography, and fuel), fire weather condition (i.e., CFWIS indices) was the most influential group in 5 of the 10 ecozones [TC, TSE, BP, HP, and boreal cordillera (BC)], with three of them (TC, TSE, and BP) displaying a dominant (>50%) influence (Fig. 2, B and C). Fuel characteristics played a major role in the remaining five ecozones, most strongly in boreal shield east (BSE) and montane cordillera (MC). No foremost influence was found for topographic variables; however, this group did show significant contribution in TSW and taiga plains (TP).

### Fire burns more severely in the north especially in the summer months

On average, ~48 (±16) days per year were conducive to high burn severity across Canadian forests during 1981 to 2020 (Fig. 3A; figs. S2 and S3 for low– and moderate–burn severity classes; and figs. S4 to S6 for maps of 75th percentile, 95th percentile, and maximum days). HP, TSE, and TSW had the greatest annual number of high–burn severity days (56 to 82 days) followed by TP and BSE (48 to 49 days) (table S5). Overall, modeled grids with most days estimated as high or moderate burn severity were primarily located in the areas dominated by coniferous forests in the northwestern uplands and northeastern region, whereas grids with many low–burn severity days were mainly in the southern broadleaf and mixed-wood forests and in the southwestern mountain forests (Fig. 1 and Fig. 3A).

The overall most severe month (i.e., the month that was most frequently identified as the most severe month in each year) occurred predominantly in summer (June to August) for most ecozones (Fig. 4A and fig. S7). TSW, TSE, HP, and TP demonstrated the greatest potential for high burn severity (56 ± 5%) for their respective overall most severe months (Fig. 4B). Some grids in southern TP, eastern BP, southern boreal shield west (BSW), and southern BSE, primarily covered by broadleaf and mixed-wood forests, were more prone to high burn severity in spring, but their high–burn severity potential was relatively lower than those prone to high burn severity in summer (Fig. 1 and Fig. 4, A and B).

In 2023, an escalated burn severity potential, but similar spatial pattern, was found across Canada compared with 1981 to 2020 (Fig. 3B). This escalation was most pronounced in the northern ecozones (TSE, TSW, HP, BC, TC, and TP), where the number of high–burn severity days in 56.6% (±9.9%) of the modeled grids exceeded their 40-year average. This proportion decreased to 29.1% (±10.0%) in the southern regions (MC, BP, BSW, and BSE) (Fig. 3B). Notably, modeling showed that 11.5% (±11.4%) of grid points in the north experienced more than their 40-year maximum number of days conducive to high-severity burning compared with only 2.2% (±2.1%) in the south. The most severe months of 2023 showed a similar pattern to that summarized over 1981 to 2020 (fig. S8).

### Canadian forests are more conducive to high-severity fires in the past two decades

Trend analysis showed that across all ecozones, ~6% (0.54 to 14.64%) of the modeled grids

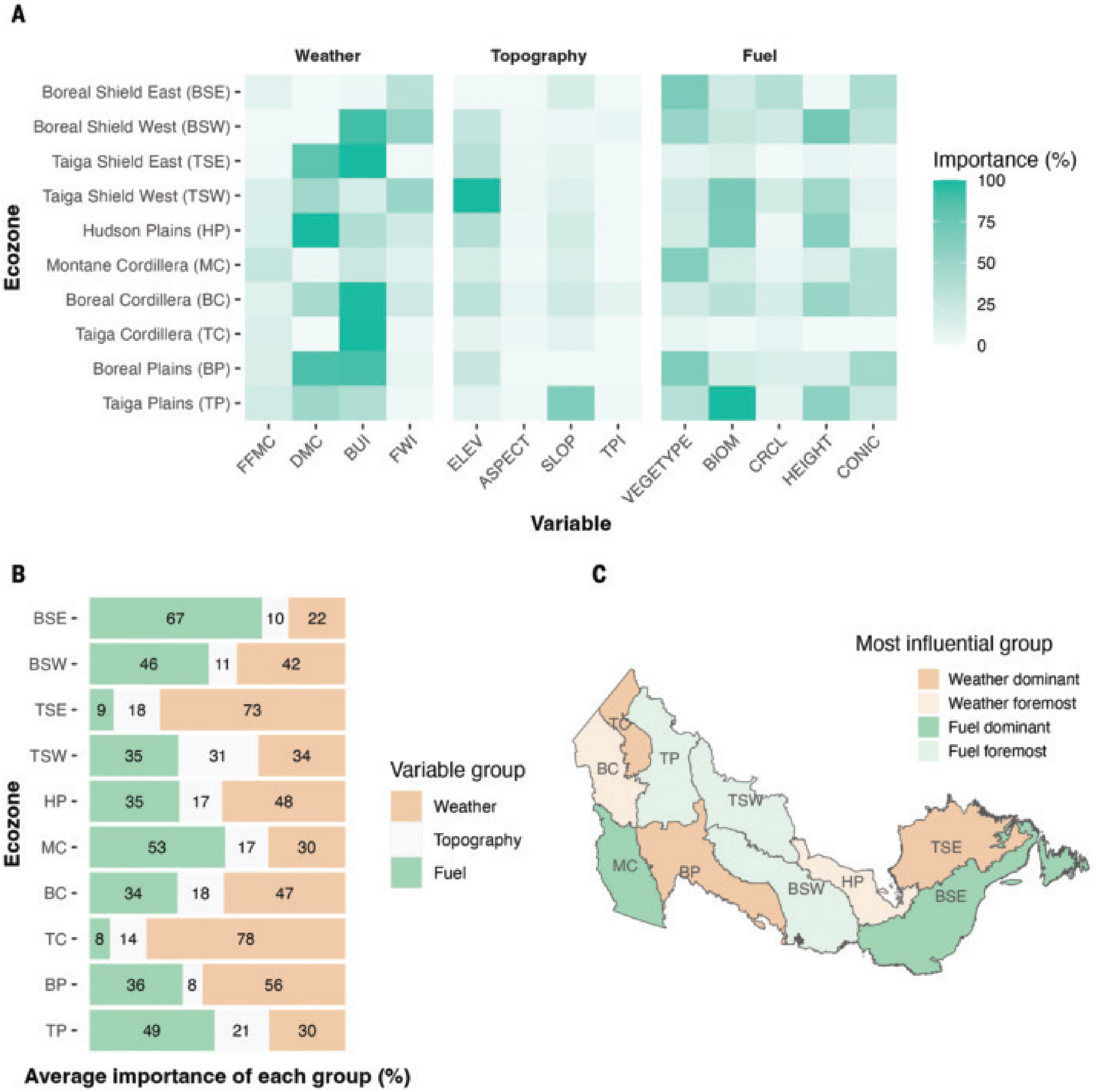
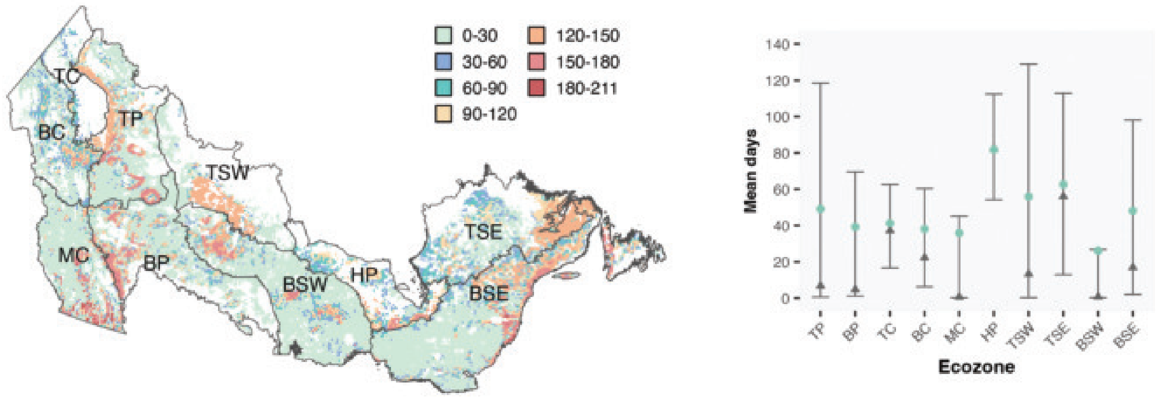


**Fig. 2. Influence of different environmental drivers on burn severity in the ecozones.** (**A** to **C**) Importance of each individual predictor (A), average importance of each group of predictors (B), and most influential group of predictors in each ecozone (C). In (C), "foremost" describes the group with the highest average importance in the ecozone, and a group is considered "dominant" if it has an average importance ≥50% (see materials and methods for modeling procedure and validation). FFMC, fine fuel moisture code; FWI, fire weather index; ELEV, elevation; SLOP, slope; TPI, topographic position index; VEGETYPE, vegetation type; BIOM, biomass; CRCL, crown closure; CONIC, coniferous species cover percentage.

exhibited significant ($\alpha_{\text{global}}$ = 0.05) increases in the number of high–burn severity days during 1981 to 2020, with a mean rate of +0.5 days per year (Fig. 5 and table S6). TSE (14.64%) and TP (12.82%) had the greatest increases, followed by TC (9.23%) and BC (7.17%). Most of the increases in high–burn severity days, particularly in TP, TC, and BC, correspond with significant increases in fire-conducive weather conditions (fig. S9). Seasonally, increases in high–burn severity days primarily occurred in spring and autumn (figs. S10 to S13), coinciding with areas that had the overall most severe months in summer (Fig. 4). Conversely, significant decreases were very limited (1.6 ± 1.5%), with the highest proportion (4.5%) occurring in BSE over the 40-year period. The number of days conducive to low- and moderate-severity burning (figs. S14 to S17) showed significant changes mainly in TC, BC, TSE, TP, and MC, with increases in low–burn severity days primarily estimated in the southern areas of these ecozones, which burned most severely in spring or autumn (Fig. 4A). Moreover, few grid points were found with significant trends in the annual most severe potential (potential for high burn severity for the most severe month in each year; table S3) (figs. S18 and S19), and TC was the only ecozone with a high proportion (10%) of increases during 1981 to 2020.

Notably, significant increases in the number of high–burn severity days were usually found in the most recent 20 years (2001 to 2020) compared

**A** Mean annual number of days conducive to high-severity burning during 1981-2020



**B** 2023 anomaly in comparison to the last 40 years

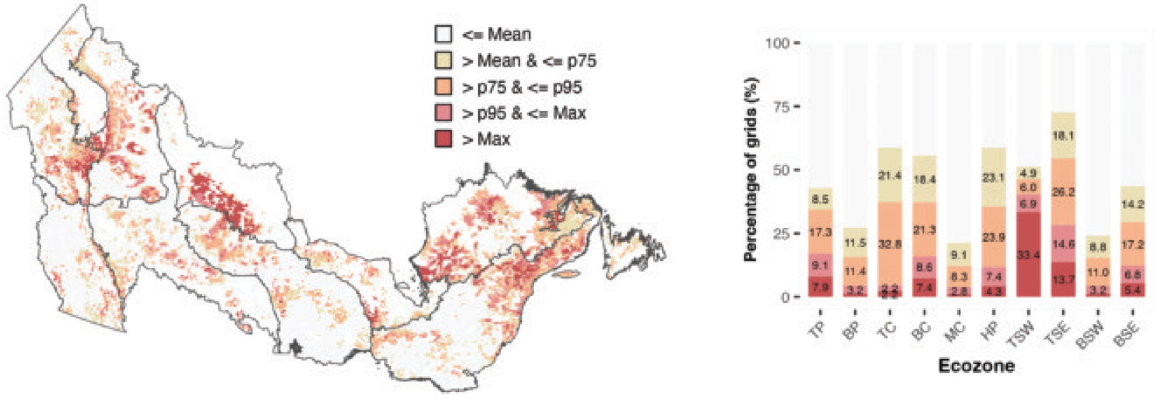


**Fig. 3. Spatial pattern of high–burn severity days and 2023 anomaly.** (**A**) Map (left) and statistical summary (right) of mean annual number of days conducive to high-severity burning over the 40 years (1981 to 2020). Green dots and gray triangles are the means and medians, respectively, and vertical lines represent the range of 25th to 75th percentiles (see table S5 for the summary statistics). (**B**) Difference in estimated high–burn severity days in 2023 compared with the 40-year mean, 75th percentile (p75), 95th percentile (p95), and maximum values for each modeled grid (left) as well as the statistical summary by ecozone (right). Bars without displayed percentages have values <2%.

with the preceding 20 years (1981 to 2000) (Fig. 5). During 2001 to 2020, BC had the largest proportion (10.5%) of modeled grid points with significant increases in high–burn severity days (rate of +1.8 days per year), followed by TC and TP, with a proportion of ~8.5% (rate of +1.5 days per year). By contrast, all ecozones showed limited (<2.1%) grid points with significant increases in the number of high–burn severity days during the preceding 20 years. Overall, the mean number of high–burn severity days in the most recent 20 years was significantly (α = 0.001) greater by 0.4 to 5.0 days compared with the preceding 20 years across the ecozones (table S7). Although about half of the ecozones showed significant changes in the mean annual most severe potential (table S3), only TC and HP demonstrated significant and notable increases—3.9% and 1.7%, respectively—in the most recent 20 years compared with the prior 20 years (table S8).

## Discussion

This study underscores the complex interplay of factors that influence burn severity in Canadian forests, focusing on spatial pattern and trends over a 40-year period and emphasizing the pressing need for proactive strategies to mitigate the increasing threat posed by climate change. We found that two fuel aridity measures, BUI and DMC, were the most influential drivers of burn severity. BUI and DMC were designed to capture the potential amount of fuel available for combustion and to track the condition of the duff layer fuel moisture (*33*, *34*). Our results demonstrate the critical role that drought plays in burn severity of the Canadian forests. Given the fundamental function of the CFWIS in Canada, this finding can contribute to fire management practices, such as assessing fire risk and planning prescribed fires. The connection between these measures and fire intensity (*33*, *34*, *42*) also indicates that increasing intensity may lead to heightened burn severity (*43*).

Additionally, our findings demonstrate the important influence of vegetation characteristics in certain regions, aligning with previous studies (*15*, *29*, *30*). Fuel availability (biomass and height) played an important role in driving the severity of fires that burned in the eastern TP and western TSW, where the variability of weather conditions over time was relatively low (*44*, *45*). Prefire vegetation types were found to be the key driver in areas with vegetative diversity, including the mountain and subalpine forests in MC and the temperate broadleaf and mixed-wood forests in BSE and BSW. By contrast, topographic variables showed no foremost influence, which could be attributed to their close correlations with vegetation distribution and local weather conditions as well as the limited topographic variation throughout most of our study area. Our previous study

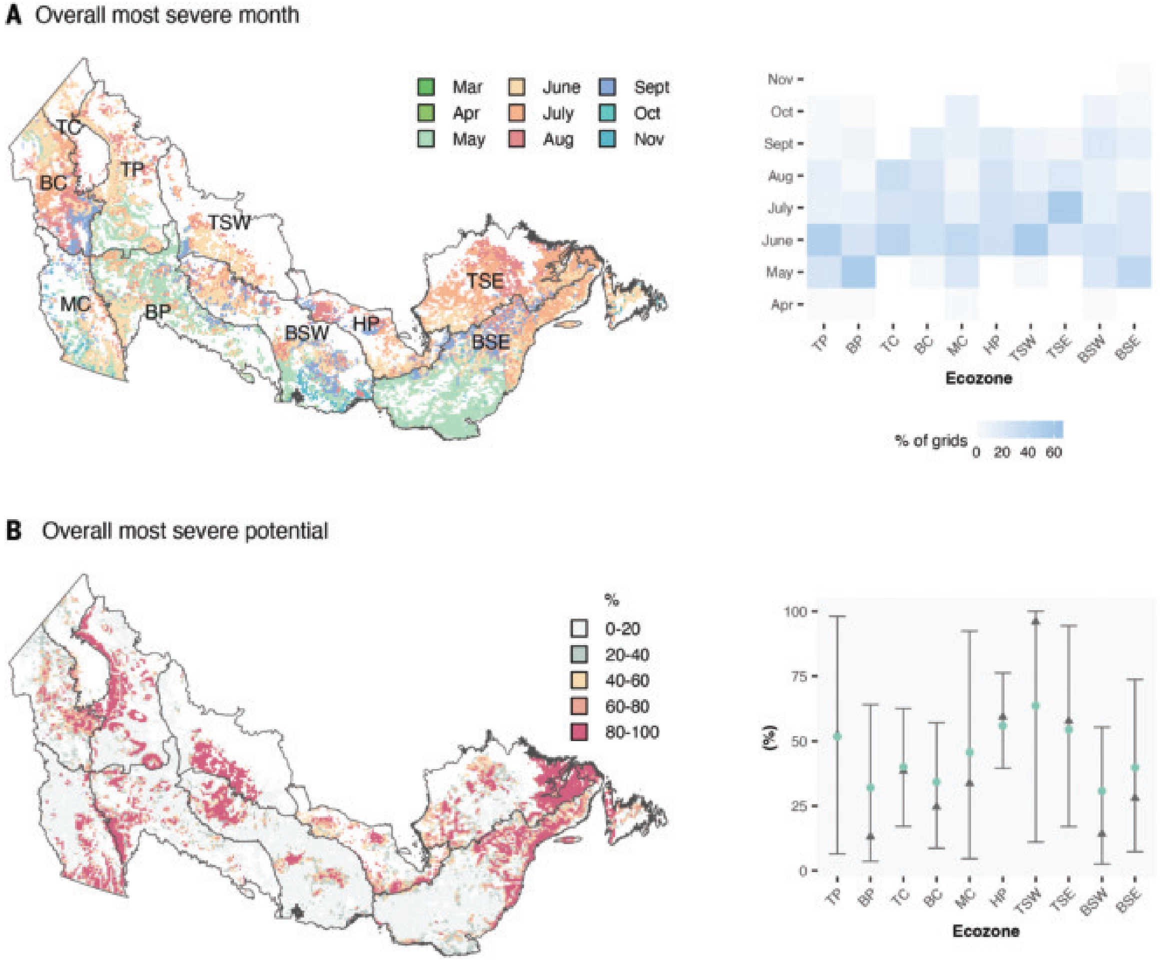


**Fig. 4. Spatial and seasonal pattern of the most severe condition.** (**A** and **B**) Maps and summaries of the overall most severe months (A) and the overall most severe potential (B) during 1981 to 2020 within the study area; green dots and gray triangles are the means and medians, respectively, and vertical lines represent the range of 25th to 75th percentiles (see table S5 for the summary statistics). The overall most severe month represents the month that was most frequently identified as the most severe month in each year during 1981 to 2020, and its mean monthly potential for high burn severity over 40 years is referred to as the overall most severe potential (table S3). Here, monthly potential for high burn severity refers to the proportion of days per month that were estimated to burn at high severity, and the most severe month in each year is the month with the largest monthly potential for high burn severity. A breakdown of the most severe month in each year by ecozone can be found in fig. S7.

in the Rocky Mountain region (*30*) showed considerable influences from topographic variables, especially elevation, on burn severity. This may indicate that the scale of analyses is a factor in the discrepancy.

High–burn severity days were most frequent during the summer months in the northwestern uplands and northeastern region, where there are extensive coniferous trees. Coniferous forests are generally susceptible to severe fires owing to flammable characteristics attributed to their needlelike leaves, higher resin content, and ladderlike branch structure (*42*, *46*). Additionally, summer sees the highest occurrence of lightning-caused fires in these areas; these fires generally produce vast burned areas with high burn severity (*28*, *47*, *48*). By contrast, estimated low–burn severity days were primarily in southern areas, where their most severe months are in spring or autumn—a period when weather conditions are mild and low–burn severity, human-caused fires typically occur (*28*, *47*, *48*). The greatest potential for high burn severity in spring in these areas could also be attributable to broadleaf tree phenology and the so-called "spring dip" phenomenon in conifers (*49*).

Canada has experienced escalating wildfire seasons over recent decades (*1*, *2*, *32*), and burn severity is expected to increase accordingly. Supporting this hypothesis, we found that all ecozones exhibited significant increases in the number of high–burn severity days during 1981 to 2020, with an average of 6% of the modeled area showing this trend. The northeast and northwest ecozones showed the greatest increases, ranging between 7 and 15%, which were expected to exacerbate regional fire season because these areas had a severe burning condition in summer and the increases primarily occurred in spring and autumn. These results are attributed to daily weather variability, the only dynamic inputs in our modeling exercise. The increases were associated with more fire-conducive days and partially coincide with areas that experienced increased extreme fire weather across Canada (*44*, *45*). The widespread increases, along with limited decreases, in high–burn severity days during 1981 to 2020 indicate the increasingly severe fire situation and more challenging fire season under the changing climate in Canada.

By dividing the 40-year period into two 20-year segments, 1981 to 2000 and 2001 to 2020, we found more-severe conditions in the most recent 20 years. The year 2000 has been identified as a global turning point in climate change, with increasingly extreme weather events having been observed since then (*45*, *50*, *51*). The pre-2000 period has been used as a baseline

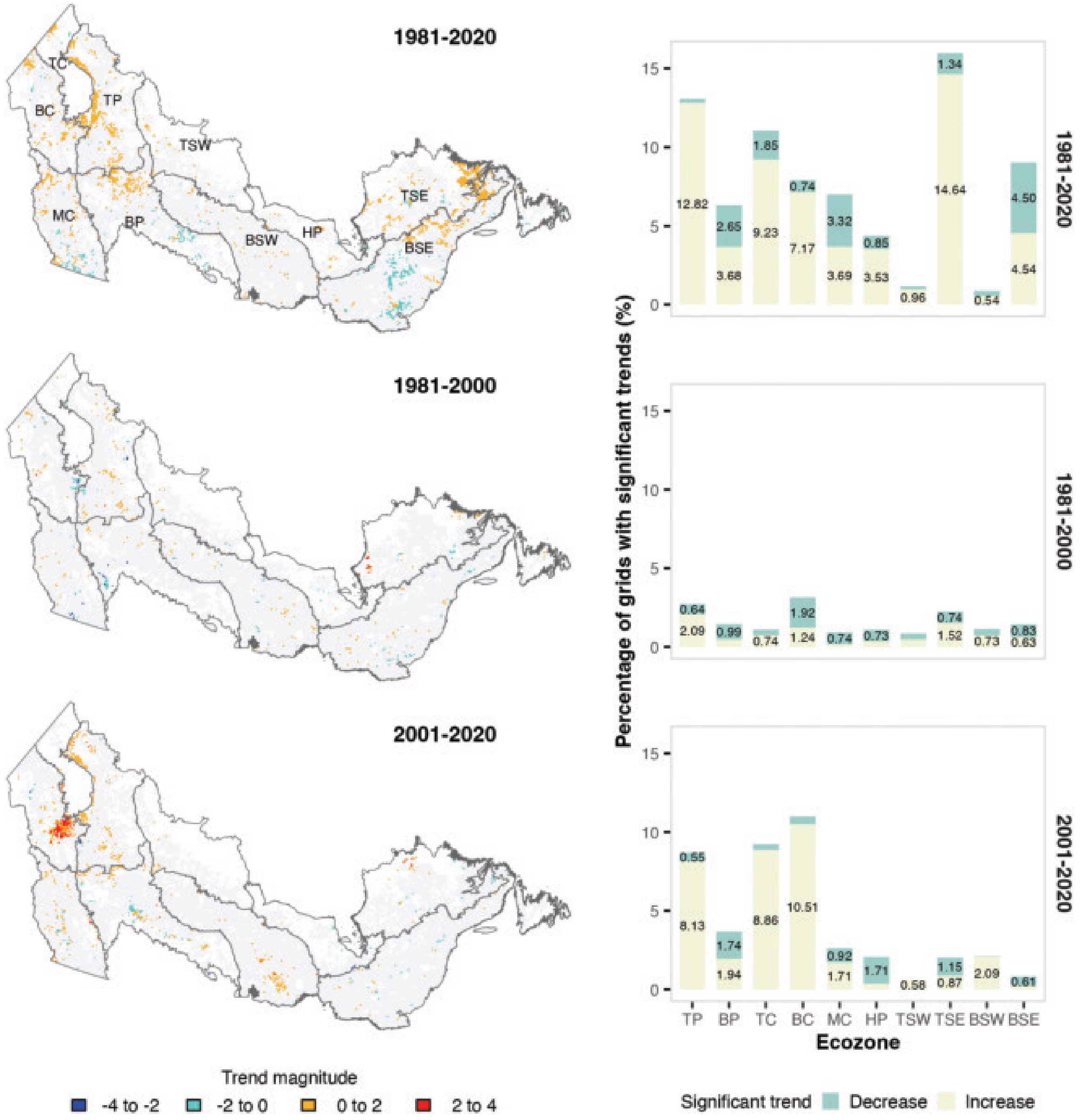


**Fig. 5. Temporal changes of high–burn severity potential during 1981 to 2020.** Significant trends in the annual number of days conducive to high-severity burning during 1981 to 2020 (top), 1981 to 2000 (middle), and 2001 to 2020 (bottom) and percentage of grids with significant trends during each period in the ecozones. Gray color in the maps represents an insignificant trend at a significant level $\alpha_{\mathrm{global}}$ = 0.05. Bars without percentages have values <0.5%.

to assess climate changes and their associated impacts. Many aspects of fire activity, such as fire season length, extreme fire weather events, fire intensity, and annual burned area, have been estimated to be more pronounced in the 21st century, compared with the pre-2000 baseline, across North America (*1*, *2*, *43*, *50*). Our results demonstrate the heightened fire effects, measured by burn severity, under recently changing climates in Canada. They complement previous observations of intensified burn severity and increasingly large high-severity areas in the western US, which have been attributed to warmer and drier climatic conditions (*24*, *26*). The 2023 estimates demonstrate widespread escalation in high–burn severity days compared with the past 40 years. In this extreme fire year, more severe conditions were found in the northwest and northeast areas, corresponding to the record-breaking large fires that occurred in British Columbia, Alberta, the Northwest Territories, and Quebec. Our findings indicate that more severe fire effects are expected in Canada because of the warming climate, with the most pronounced effects occurring in northern high latitudes; this affirms the findings of previous studies (*1*, *31*).

Maps of burn severity estimates have crucial implications for fire management and preparedness. Some areas in BSE, BSW, and BP, where we estimated relatively higher likelihood of severe burns, overlap with greater population densities and higher fire deficits (*12*, *52*). This indicates an elevated threat to local communities, warranting greater attention and concern. Maps of average high–burn severity days and the most severe fire months across Canada may also contribute to prescribed burn planning (*6*, *12*). The 40-year daily estimates of burn severity based on a static contemporary vegetation scenario enabled us to quantify the effects of climate change on fire. Ground-truth estimates of burn severity could be improved with the following efforts. First, incorporating dynamic vegetation characteristics into the models could yield better outcomes in real-time burn severity. Second, incorporating forecasts on fire occurrence and spread would enrich the information in actual fire effects because the estimates were made under the assumption of fire occurring at each modeled grid. Third, including field measurements across different regions of Canada could assist with the categorization of burn severity classes and contribute to the validation of model outputs.

## REFERENCES AND NOTES

1. M. Flannigan *et al*., *For. Ecol. Manage.* **294**, 54–61 (2013).
2. X. Wang *et al*., *Clim. Change* **130**, 573–586 (2015).
3. X. Wang, T. Swystun, M. D. Flannigan, *Sci. Total Environ.* **830**, 154752 (2022).
4. B. Zheng *et al*., *Science* **379**, 912–917 (2023).
5. P. Jain *et al*., *Nat. Commun.* **15**, 6764 (2024).
6. C. H. Key, N. C. Benson, "Landscape assessment (LA)," in "FIREMON: Fire Effects Monitoring and Inventory System," D. C. Lutes, Ed. (US Department of Agriculture, Forest Service, Rocky Mountain Research Station, general technical report RMRS-GTR-164-CD164, 2006).
7. J. E. Keeley, *Int. J. Wildland Fire* **18**, 116–126 (2009).
8. K. C. Hannah, J. E. Put, J. Shonfield, *For. Ecol. Manage.* **528**, 120632 (2023).
9. K. T. Davis *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2208120120 (2023).
10. M. G. Turner, W. W. Hargrove, R. H. Gardner, W. H. Romme, *J. Veg. Sci.* **5**, 731–742 (1994).
11. M. G. Turner, K. H. Braziunas, W. D. Hansen, B. J. Harvey, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11319–11328 (2019).
12. L. M. Johnston *et al*., *Environ. Rev.* **28**, 164–186 (2020).
13. S. L. Stephens *et al*., *Science* **342**, 41–42 (2013).
14. T. V. Loboda, N. H. F. French, C. Hight-Harf, L. Jenkins, M. E. Miller, *Remote Sens. Environ.* **134**, 194–209 (2013).
15. E. Whitman *et al*., *Ecosphere* **9**, e02128 (2018).
16. J. D. Miller, A. E. Thode, *Remote Sens. Environ.* **109**, 66–80 (2007).
17. S. Parks, G. Dillon, C. Miller, *Remote Sens.* **6**, 1827–1844 (2014).
18. R. J. Hall *et al*., *Int. J. Wildland Fire* **17**, 476–489 (2008).
19. N. O. Soverel, N. C. Coops, D. D. B. Perrakis, L. D. Daniels, S. E. Gergel, *Int. J. Wildland Fire* **20**, 518–531 (2011).
20. E. Whitman, M.-A. Parisien, L. M. Holsinger, J. Park, S. A. Parks, *Int. J. Wildland Fire* **29**, 995–1008 (2020).
21. C. W. Miller, B. J. Harvey, V. R. Kane, L. M. Moskal, E. Alvarado, *Int. J. Wildland Fire* **32**, 449–475 (2023).
22. L. Breiman, *Mach. Learn.* **45**, 5–32 (2001).
23. L. Collins, P. Griffioen, G. Newell, A. Mellor, *Remote Sens. Environ.* **216**, 374–384 (2018).
24. Y. Huang, Y. Jin, M. W. Schwartz, J. H. Thorne, *Environ. Res. Lett.* **15**, 104033 (2020).
25. V. R. Kane *et al*., *For. Ecol. Manage.* **358**, 62–79 (2015).
26. S. A. Parks, J. T. Abatzoglou, *Geophys. Res. Lett.* **47**, e2020GL089858 (2020).
27. T. Woolley, D. C. Shaw, L. M. Ganio, S. Fitzgerald, *Int. J. Wildland Fire* **21**, 1–35 (2012).
28. L. Guindon *et al*., *Can. J. For. Res.* **51**, 1230–1244 (2021).
29. X. J. Walker *et al*., *Nat. Clim. Chang.* **10**, 1130–1136 (2020).
30. W. Wang *et al*., *Geophys. Res. Lett.* **49**, e2022GL097945 (2022).
31. M. D. Flannigan, B. D. Amiro, K. A. Logan, B. J. Stocks, B. M. Wotton, *Mitig. Adapt. Strategies Glob. Change* **11**, 847–859 (2006).
32. M. D. Flannigan, M. A. Krawchuk, W. J. De Groot, B. M. Wotton, L. M. Gowman, *Int. J. Wildland Fire* **18**, 483 (2009).
33. B. D. Lawson, O. B. Armitage, "Weather Guide for the Canadian Forest Fire Danger Rating System" (Canadian Forest Service, Northern Forestry Centre, 2008).
34. C. E. van Wagner, "Development and Structure of the Canadian Forest Fire Weather Index System" (Canadian Forestry Service, technical report 35, 1987).
35. S. A. Parks, *Int. J. Wildland Fire* **23**, 215–223 (2014).
36. X. Wang *et al*., *Glob. Change Biol.* **20**, 2518–2530 (2014).
37. X. Wang *et al*., *Sci. Total Environ.* **869**, 161831 (2023).
38. M. G. Kendall, *Rank Correlation Methods* (Griffin, 1948).
39. H. B. Mann, *Econometrica* **13**, 245–259 (1945).
40. P. K. Sen, *J. Am. Stat. Assoc.* **63**, 1379–1389 (1968).
41. H. Theil, *Indag. Math.* **12**, 173 (1950).
42. Forestry Canada Fire Danger Group, "Development and Structure of the Canadian Forest Fire Behavior Prediction System" (Forestry Canada, information report ST-X-3, 1992).
43. B. M. Wotton, M. D. Flannigan, G. A. Marshall, *Environ. Res. Lett.* **12**, 095003 (2017).
44. P. Jain, X. Wang, M. D. Flannigan, *Int. J. Wildland Fire* **26**, 1009–1020 (2017).
45. P. Jain, D. Castellanos-Acuna, S. C. P. Coogan, J. T. Abatzoglou, M. D. Flannigan, *Nat. Clim. Chang.* **12**, 63–70 (2022).
46. M. E. Alexander, M. G. Cruz, "Crown fire dynamics in conifer forests" in "Synthesis of Knowledge of Extreme Fire Behavior" (US Department of Agriculture, Forest Service, Pacific Northwest Research Station, general technical report PNW-GTR-854, 2011), pp. 107–142.
47. C. C. Hanes *et al*., *Can. J. For. Res.* **49**, 256–269 (2019).
48. B. J. Stocks *et al*., *J. Geophys. Res.* **107**, FFR 5-1–FFR 5-12 (2002).
49. M. A. Parisien, Q. E. Barber, M. D. Flannigan, P. Jain, *Glob. Change Biol.* **29**, 6106–6119 (2023).
50. M. W. Jones *et al*., *Rev. Geophys.* **60**, e2020RG000726 (2022).
51. A. P. Williams *et al*., *Science* **368**, 314–318 (2020).
52. M.-A. Parisien *et al*., *Nat. Commun.* **11**, 2121 (2020).
53. Copernicus Climate Change Service, Complete ERA5 global atmospheric reanalysis (2023); https://doi.org/10.24381/cds.143582cf.
54. L. Guindon *et al*., SCANFI: The Spatialized CAnadian National Forest Inventory data product (Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, 2023); https://doi.org/10.23687/18e6a919-53fd-41ce-b4e2-44a9707c52dc.
55. W. Wang, Modelling burn severity in Canadian forests (v1.0.0), version 1.0.0, Zenodo (2024); https://doi.org/10.5281/zenodo.11999873.

## ACKNOWLEDGMENTS

The authors thank N. C. Coops and C. Mulverhill from the Integrated Remote Sensing Studio, Faculty of Forestry, the University of British Columbia, for insightful comments on the manuscript. The authors also thank J. Oliver from the Great Lakes Forestry Centre, Canadian Forest Service, and K. Levesque from the Northern Forestry Centre, Canadian Forest Service, for their help with writing. **Funding:** This work was funded by the APFNet-UBC National Park Research Project (APFNet 2017 Sp2-UBC). **Author contributions:** Conceptualization: X.W., W.Wa., M.D.F., and G.W. Methodology: W.Wa., X.W., L.G., T.S., and D.C.-A. Investigation: W.Wa., X.W., M.D.F., and G.W. Visualization: W.Wa. Supervision: G.W. and X.W. Writing – original draft: W.Wa. Writing – review and editing: W.Wa., X.W., M.D.F., T.S., L.G., W.Wu, and G.W. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The National Burned Area Composite (NBAC) dataset and the MODIS and VIIRS hotspots used to derive daily fire growth data from 2001 to 2020 are available from https://cwfis.cfs.nrcan.gc.ca/datamart and https://firms.modaps.eosdis.nasa.gov/active_fire/, respectively. The Canada Landsat Burned Severity product (CanLaBS) used to derive the burn severity data is available from https://open.canada.ca/data/en/dataset/b1f61b7e-4ba6-4244-bc79-c1174f2f92cd. The ERA5 data used to derive the gridded CFWIS indices for this study are available at (*53*). The digital elevation model (DEM) from the 3 arc sec (~90 m) EarthEnv-DEM90 product is available from http://www.earthenv.org/DEM.html. The Spatialized Canadian National Forest Inventory (SCANFI) data product used to derive fuel characteristics is available at (*54*). No custom mathematical algorithms were used in our analyses. The R code is available at (*55*).

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ado1006
Materials and Methods
Supplementary Text
Figs. S1 to S20
Tables S1 to S13
References (*56*–*89*)

Submitted 17 January 2024; accepted 28 October 2024
10.1126/science.ado1006

---

MIGRATION

# Bats surf storm fronts during spring migration

**Edward Hurme[1,2,3]*, Ivan Lenzi[1], Martin Wikelski[1,2,3], Timm A. Wild[1]†, Dina K. N. Dechmann[1,2,3]†**

**Long-distance migration, common in passerine birds, is rare and poorly studied in bats. Piloting a 1.2-gram IoT (Internet of Things) tag with onboard processing, we tracked the daily location, temperature, and activity of female common noctules (*Nyctalus noctula*) during spring migration across central Europe up to 1116 kilometers. Over 3 years, 71 bats migrated tens to hundreds of kilometers per night, predominantly with incoming warm fronts, which provided them with wind support. Bats also showed unexpected flexibility in their ability to migrate across a wide range of conditions if needed. However, females leaving toward the end of the season showed higher total activity per distance traveled, a possible cost for their flexible migration timing.**

Billions of birds migrate each year, using a set of favorable conditions to inform migration decisions. In the more cryptic and nocturnal bats, long-distance migration occurs in only a few, both insectivorous and frugivorous species in temperate and tropical regions, yet little is known about their migration routes or strategies (*1*). Long-distance migrants face not only high energetic costs but also multiple threats such as wind power turbines, decreasing insect abundance, and multiple effects of climate change (*2*–*4*). Rising temperatures from global warming may cause earlier spring departures that may be beneficial or lead to a mismatch between arrival timing and food abundance (*5*). Emerging evidence indicates that populations of migratory bat species are shifting ranges and/or decreasing range size, but the causes for this remain poorly understood (*6*).

Migrants must continually decide whether to proceed to the next leg. Departure cues appear to be linked to local weather conditions, in particular, favorable wind conditions (*7*–*9*). This is supported by the observation that some bat species benefit from wind support during foraging flights (*10*, *11*) and the small

---

[1]Department of Migration, Max Planck Institute of Animal Behavior, Radolfzell, Germany. [2]Centre for the Advanced Study of Collective Behaviour, University of Konstanz, Konstanz, Germany. [3]Department of Biology, University of Konstanz, Konstanz, Germany.
***Corresponding author. Email: edwardhurme@gmail.com**
†These authors contributed equally to this work.

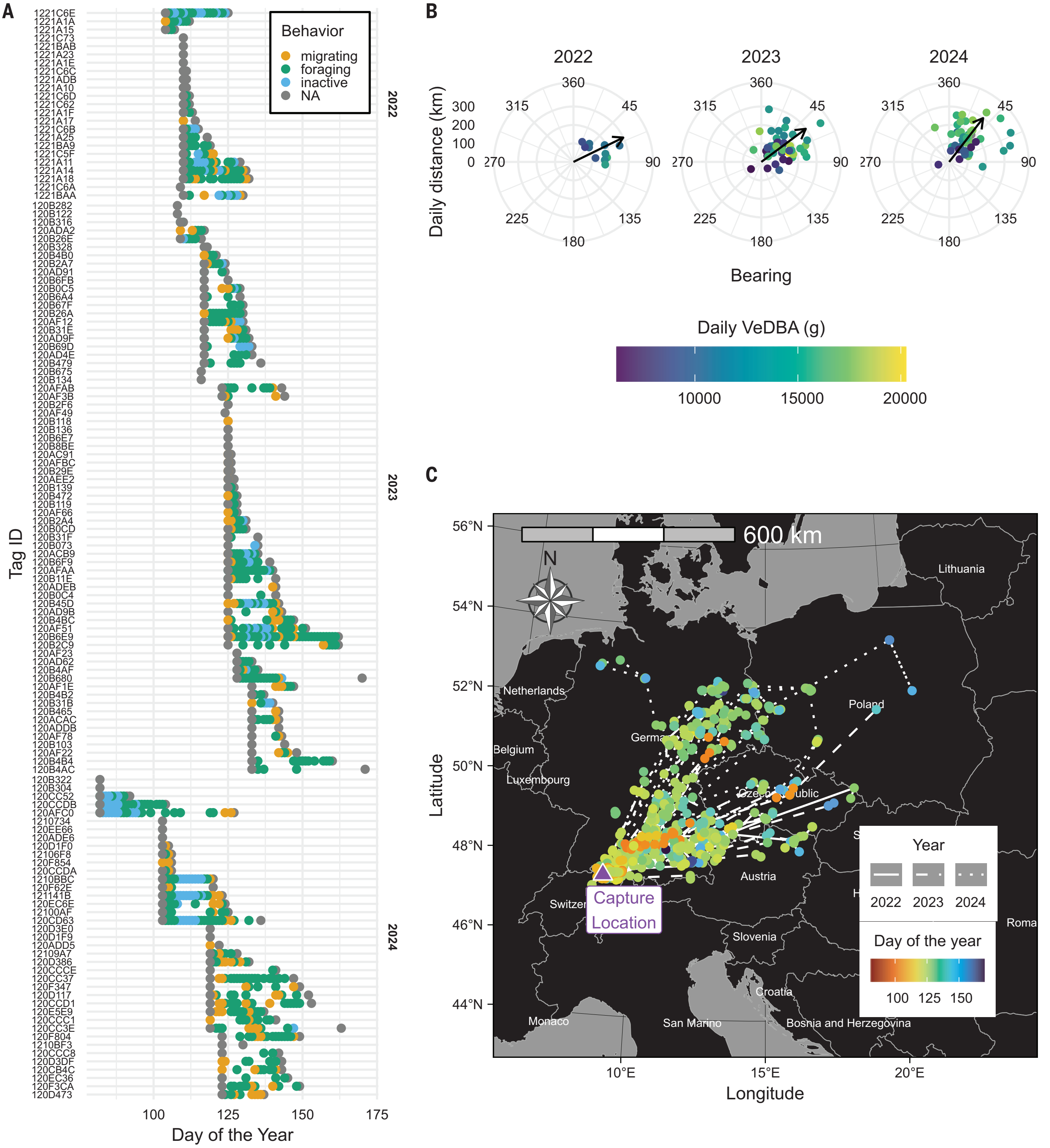


**Fig. 1. Female common noctule spring migration.** (**A**) Assigned daily behavior of each tagged bat from evaluations of daily distance and daily VeDBA, a measure of activity. Inactive and foraging (blue and green, respectively) were determined from a threshold of daily total VeDBA, and migration (orange) was estimated from a daily distance threshold. NA (not available) values correspond to locations with missing values for distance or VeDBA, typically representing the day of capture. Each tag ID is associated with a single individual and is ordered chronologically by deployment time. (**B**) Direction and distance of each migration step for 2022, 2023, and 2024 are colored by the daily VeDBA. The average migration bearing for each year is indicated with an arrow (2022, 63.9°; 2023, 53.3°; 2024, 37.8°). (**C**) Map of Central Europe showing the individual tracks, connected by solid lines in 2022, dashed lines in 2023, and dotted lines in 2024; each daily point is colored by day of the year. The capture location is shown by a purple triangle in the northeast region of Switzerland.

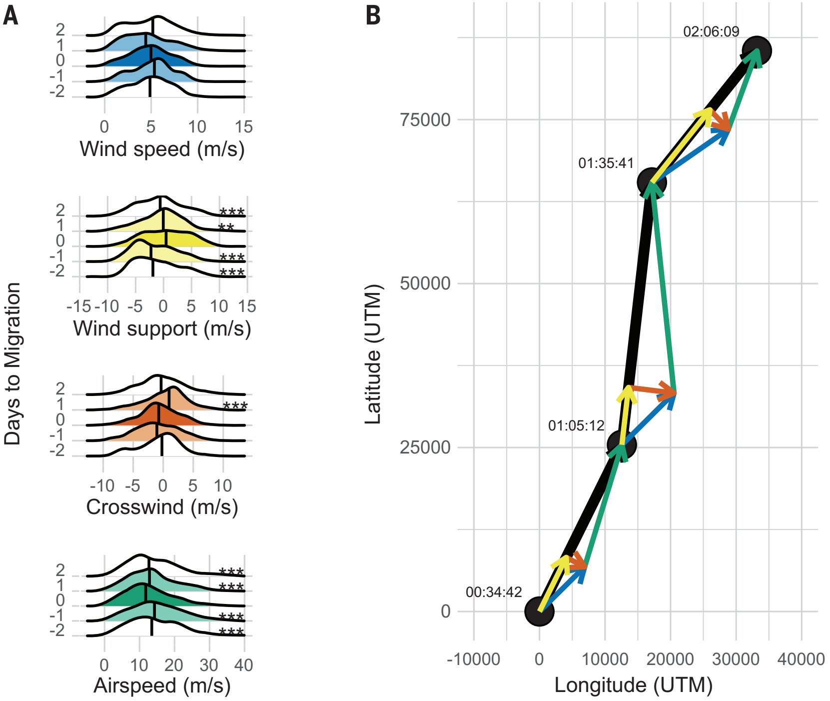


**Fig. 2. Bats often migrate on nights with higher wind support.** (**A**) Flight conditions during migration flights compared with 2 days before and after migration, showing density plots of wind speed (blue), wind support (yellow), crosswind speeds (orange), and airspeed (green). Vertical bars in the density plots show the median values for each group. Asterisks indicate significant paired Wilcoxon tests comparing instantaneous wing conditions during actual migration flights with adjacent nights (****P* < 0.001, ***P* < 0.01). (**B**) Schematic representation of calculated wind measurements on a segment of the migration. Black points show the observed locations of a bat in a migratory flight (tag ID: 120ACB9), with black lines connecting each point. Blue arrows represent the instantaneous wind speed and direction at that observed location. Yellow and orange arrows estimate the wind support and crosswind speeds, respectively, encountered by the bat at this time. The green arrows correspond to the airspeed at which the bat would need to fly to generate the observed ground speed under these wind conditions.

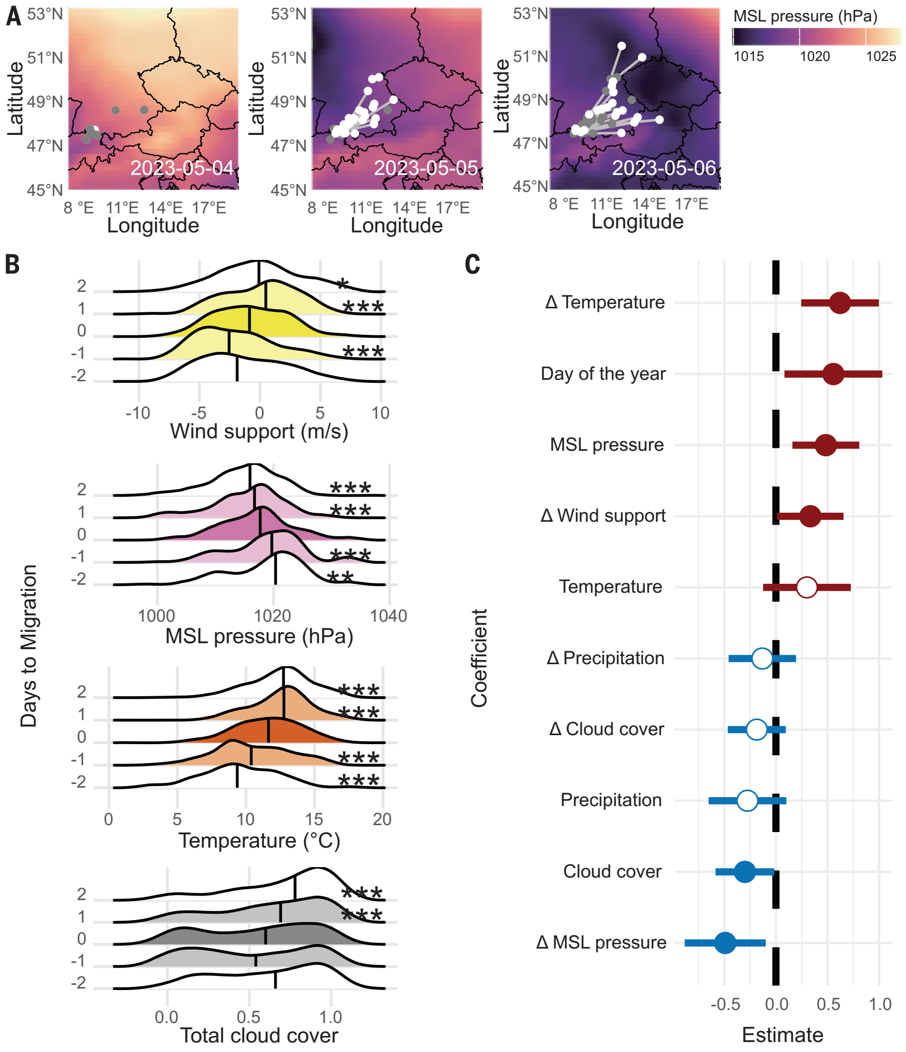


**Fig. 3. Bats migrate before storm fronts.** (**A**) Locations of bats over three nights as a front moves across Europe. Air pressure data [ECMWF ERA5 hourly mean sea level (MSL) pressure at 0:00 UTC] illustrate the front's northeastward movement. Bat paths are shown for 3 days, with locations connected by gray lines. Points represent bat locations at 0 and 1 day before the labeled timestamp, filled in white and gray, respectively. (**B**) Average nightly weather conditions 2 days before and after migrations, showing wind support (yellow), MSL pressure (violet), temperature (dark orange), and cloud cover (gray). Vertical bars in the density plots show the median values for each group. Asterisks indicate significant paired Wilcoxon tests comparing nightly weather conditions during actual migrations with adjacent nights (****P* < 0.001, ***P* < 0.01, **P* < 0.05). (**C**) Model coefficients of the binomial generalized linear mixed model are shown by circles colored red and blue for positive and negative effects on migration, respectively, and lines show the 95% CI. Closed circles indicate a significant effect of this coefficient (*P* < 0.05), and empty circles indicate that the 95% CI overlaps with 0. Migration is associated with high air pressure that is decreasing, later dates, and increases in temperature and wind support.

portions of their migrations that have been tracked to date (*12*). Describing the interplay between energy conservation and risking the next migration step is crucial for the understanding of this increasingly threatened phenomenon. However, owing to limitations of tracking technology, full bat migrations have not been monitored.

Our study species, the common noctule (*Nyctalus noctula*), is a wide-ranging European insectivorous bat. Predominantly female noctule bats migrate northeast each spring, covering distances of up to 1600 km from their hibernacula to maternity roosts (*13*). However, migration patterns are changing, with some local populations in the traditional wintering areas declining and wintering areas expanding farther north (*6*, *14*). At Lake Constance in southern Germany, bats often departed with wind support and high air pressure to complete relatively short migration steps interrupted by stopovers (*7*, *12*, *15*). They then used highly variable flight heights to fly across the landscape (*12*). We used the ICARUS TinyFox-Batt bio-telemetry tags (1.19 ± 0.08 g), featuring onboard sensor evaluation and data compression (edge computing), to allow data retrieval via the long-range Sigfox 0G network (https://www.sigfox.com/coverage/), to follow females and monitor departure decisions. Furthermore, tags estimate cumulative VeDBA (vectorial dynamic body acceleration, a standard measure of total acceleration calculated as the square root of the summed vector measures from an accelerometer), an acceleration-derived proxy for energy consumption (*16*), during each migration step over several weeks as bats migrated away northward from the Alps across central Europe in spring (*17*).

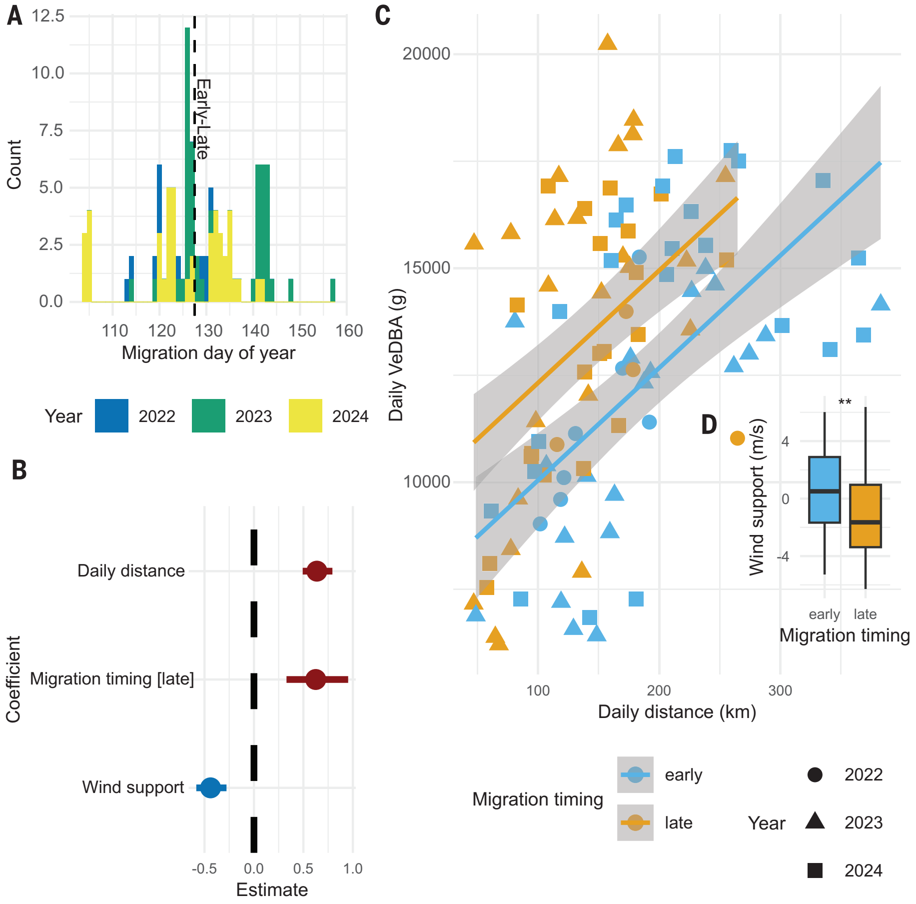


**Fig. 4. Late migrations have higher acceleration per distance traveled.** (**A**) The distribution of migration nights, colored by year, with the median migration day (7 May), here the Early-Late threshold for migration, shown as a vertical dashed line. (**B**) Model coefficients of a binomial generalized linear mixed model of daily VeDBA are shown by circles colored red and blue for positive and negative effects on energy expenditure, or VeDBA, respectively, and lines show the 95% CI. Circles indicate a significant effect of this coefficient ($P < 0.05$). Distance and late migrations are associated with higher VeDBA, whereas wind support had a negative effect on VeDBA. (**C**) Scatterplot of the daily VeDBA and daily distance for migrating bats colored by early (light blue) and late (orange) migrations. Late migrations have a higher $y$-intercept than early migrations, suggesting a general increase in energy expenditure with late migration timings. (**D**) Boxplot of the wind support for early and late migrations, colored light blue and orange, respectively. Wind support is significantly lower in late migrations (Wilcoxon-signed rank test: $n = 98$ migrations, $W = 1431$; $P < 0.01$).

## Results

### *Movement summary*

We received data from 125 individuals from 0 to 46 days (mean ± SD: 11.7 ± 11.0 days, Fig. 1A and table S1) (*18*). Some tags may have detached prematurely, limiting data collection before bats could initiate or complete their migration. From the remotely transmitted data, we obtained estimated locations, daily body acceleration measurements (VeDBA), percent daily activity (movement above a preprogrammed threshold for flying; see materials and methods), and minimum and maximum tag temperature from all tagged bats. Migration behavior could be identified during 30-min interval bursts of location fixes during the night and from daily distances (see materials and methods and figs. S1 and S2). Daily VeDBA was used to distinguish nights (periods from solar noon to solar noon) with foraging flights from 24-hour inactivity (fig. S2). Behavioral thresholds of daily acceleration and distance revealed that foraging and migration occurred at significantly higher tag temperature and activity than inactive and stationary behaviors, respectively (fig. S2, B to E). Daily VeDBA was on average 249% higher during migrating nights compared with foraging nights, whereas inactivity was only 8% of the daily VeDBA of a foraging night (fig. S3). Overall, bats spent 141 nights [95% confidence interval (CI): 134–166 nights] migrating, 489 nights (95% CI: 478–493 nights) foraging, and 110 nights (95% CI: 99–112 nights) inactive (Fig. 1A and fig. S2). Bats frequently transitioned between foraging and migration and inactive and foraging, but rarely between migration and inactive (table S2). Stopovers, periods between migration steps, were rarely observed during migrations, often because of lack of consistent signal coverage, and were highly variable in duration (median: 4 days; range: 2 to 21 days; $n = 25$ days; fig. S4). Bats foraged on the majority of stopover nights in all but one case (mean: 89% foraging nights; range: 33 to 100% foraging nights; $n = 25$ nights).

Seventy-one individuals migrated away from the tagging site in a northeasterly direction (mean ± SD: 47.6° ± 0.4°; Rayleigh test of uniformity: $\bar{R} = 0.85$, $P < 0.001$; Fig. 1, B and C) while the tag remained attached, with no significant difference in bearing between the 3 years (table S3). On migratory nights, bats traveled up to 383 km (mean ± SD: 166 ± 74 km; Fig. 1, B and C), extending known nightly flight distances of this species by nearly 200 km and well over previous estimates of migration distances (*12*, *15*, *19*). Consistent with previous tracking studies during migration (*12*), ground speeds of bats measured from burst flight recordings were on average 12.7 ± 7.1 m/s, with a maximum speed of 43.3 m/s (fig. S1B).

### *Environmental impacts on migration*

During migration flights, bats experienced a range of wind speeds (median: 5.0 m/s; range: 0.2 to 11.1 m/s). Compared with surrounding days, migratory flights had similar wind speeds, yet bats likely had higher wind support (also known as tailwind, or the component of the wind vector that aligns with the bat's flight direction), lower cross-wind speeds, and lower airspeeds (i.e., flight speeds in relation to the column of moving air; Fig. 2 and table S4).

Bats appear to migrate before incoming pressure fronts (Fig. 3A and table S5). The average weather conditions at each departure location during migratory nights typically had higher wind support (median: 0.5 m/s; range: −8.8 to 8.4 m/s; Fig. 3B). Temperature also increased from the previous nights, whereas air pressure decreased (Fig. 3, B and C, and fig. S5). To test the influence of the environment on migratory behaviors, we fit a binomial generalized linear mixed model (GLMM) to daily migratory versus stationary behaviors for all bats. Migrations typically occurred on high–air pressure nights later in the spring season. These nights often had high flux with increasing wind support and air temperature and decreasing air pressure (Fig. 3B and table S6).

### *Timing of migration*

Individuals only migrated on average for 2.0 ± 1.3 nights (range: 1 to 5 nights, $n$ = 71 bats) while tags were attached, but migrations occurred over nearly 2 months in spring (27 days: 14 April to 11 May 2022; 48 days: 19 April to 6 June 2023; 45 days: 12 April to 27 May 2024; Fig. 1A), showing a wide range of timing for female bats. There was no relationship between capture mass or added mass percentage by the tag on departure date (table S7) or total distance traveled; however, duration of tagging was positively related to total migration distance (table S8). High–migration-activity nights were defined by the mean percentage of bats migrating per night. Nights with higher than average migration activity typically had higher air temperature and lower precipitation with decreasing cloud cover and decreasing precipitation (table S9).

To test the hypothesis that bats can time their migrations to benefit from nights with good wind support, we modeled the daily VeDBA (a linear correlate of energy expenditure), using daily distance traveled, wind support interpolated along the migration path, and whether migration steps were early or late migrations. The median migration day was 7 May or the 127th Julian date, which was used as a threshold between early and late migrations (Fig. 4A). Daily VeDBA increased significantly with distance traveled (Fig. 4, B and C). Daily VeDBA, and thus likely relative energy expenditure, decreased with increasing wind support (Fig. 4B). The timing of migration revealed that late migrations had a higher intercept and therefore overall higher daily VeDBA, which is partially explained by lower wind support during late migrations (Wilcoxon-signed rank test: $P < 0.003$; Fig. 4D).

## Discussion

We used a recently developed tag with edge computing capabilities to track the extended spring migration of female common noctules and provide evidence that bats used wind support to increase speeds and reduce energy costs. Bats predominantly migrated during nights before incoming warm fronts, potentially to avoid adverse weather and benefit from the wind support provided by the warm front. High-activity migratory nights were warm and with little precipitation, suggesting that absolute temperature is a trigger for many bats to migrate. These data will now allow comparison to better-studied migratory systems. Birds and bats show similar responses to temperature and barometric pressure during spring migrations (*20*). Generally, daily changes in parameters—i.e., increasing temperature, dropping pressure, and increasing wind support—were the best predictors of the main migration nights in our bats, more than specific values of parameters. Our results are consistent with Dechmann *et al.* (*7*) in that high wind support and air pressure and late day of the year were important predictors of migration, but the change in the weather parameters may be important indicators of weather fronts for triggering migration.

Not all migration steps occurred on these nights with incoming warm fronts, showing that bats can also flexibly migrate under less optimal conditions if needed. Wind support decreased for late migrants, which then resulted in higher cumulative acceleration per distance traveled (Fig. 4D). Although we have not calibrated the accelerometers to determine energetic expenditure, VeDBA is a commonly used proxy for energy expenditure (*16*, *21*), and higher VeDBA suggests that bats used more energy per distance traveled. Bats may have partly compensated for variable wind support by adjusting flight altitudes, which is highly variable in noctules (*12*). Altitudinal flight changes allow for great flexibility, possibly enabling bats to respond to local conditions and food availability along the migratory route while still successfully completing migration. On some occasions, bats have been observed to abort a migration step and return to the departure site, similar to nocturnally migrating songbirds, at least in one case likely due to poor flight conditions from low wind support or an approaching cold front (*12*). Another indicator for high flexibility was that departure date and total flight distance were not influenced by capture mass (tables S7 and S8), consistent with previous work showing that common noctules rapidly accumulate mass before migration (*4*). Before migrating, noctule bats likely use torpor at some point every day and feed every evening, including on migration nights just before departure (*4*, *15*).

Our results of later migrants spending larger amounts of energy show that the flexibility in migration timing has an energetic cost. The observed variability in migration timing may allow individuals to spend more time preparing but cause them to miss the optimal window for migration and arrive at the summering site late, expanding our understanding of the costs and benefits of accurate migration timing in bats. Although late migrants typically have less wind support, this does not fully explain this pattern. An increase in mass could also increase VeDBA, as an increase in mass should result in a higher wingbeat frequency (*22*). Females are generally pregnant in spring and can delay the embryo's development through torpor (*23*). We hypothesize that late migrants have more costly migrations owing to increased weight from advanced pregnancy. As these bats wait for the right migration conditions, they must either invest in their embryo while increasing their own energetic cost of flight or delay the development of the embryo, possibly affecting the pup's survival. This phenological flexibility may be key for their long-term survival and maintenance of migration, yet it is unclear whether this temporally dispersed migration allows individuals to better adapt to unpredictable weather conditions than more pulsed migrations.

The long-distance individual tracking made possible by the ICARUS TinyFoxBatt tag provides new context for bat migration (*17*). This technology enables the remote transmission of 1440 daily sensor measurements through edge computing. This allows for detailed insight into the conditions that bats face as they make their migration decisions. Previous studies with shorter durations and smaller sample sizes led to assumptions of migration directions. However, our current study expands the range of migration from Switzerland as bats flew further east (*12*, *15*). Ongoing monitoring and tracking of large numbers of individuals will aid in understanding the connectivity between populations, predict their response to ongoing climate change, and assist management of important migration corridors where the risk of collisions with wind turbines is high (*4*, *24*).

## REFERENCES AND NOTES

1. T. H. Fleming, in Encyclopedia of Animal Behavior (Elsevier, 2019), pp. 605–610.
2. C. G. Guglielmo, *J. Exp. Biol.* **221**, jeb165753 (2018).

3. V. Kubelka, B. K. Sandercock, T. Székely, R. P. Freckleton, *Trends Ecol. Evol.* **37**, 30–41 (2022).
4. L. S. Lehnert *et al.*, *PLOS ONE* **9**, e103106 (2014).
5. M. E. Visser, P. Gienapp, *Nat. Ecol. Evol.* **3**, 879–885 (2019).
6. K. A. Kravchenko, A. S. Vlaschenko, L. S. Lehnert, A. Courtiol, C. C. Voigt, *Biol. Lett.* **16**, 20200351 (2020).
7. D. K. N. Dechmann, M. Wikelski, D. Ellis-Soto, K. Safi, M. T. O'Mara, *Biol. Lett.* **13**, 20170395 (2017).
8. B. Haest, P. M. Stepanian, C. E. Wainwright, F. Liechti, S. Bauer, *Climatic drivers of (changes in) bat migration phenology at Bracken Cave* (Glob Chang Biol, 2020), 10.1111/gcb.15433.
9. J. L. Pettit, J. M. O'Keefe, *J. Mammal.* **98**, 1236–1248 (2017).
10. N. Sapir, N. Horvitz, D. K. N. Dechmann, J. Fahr, M. Wikelski, *Proc. Biol. Sci.* **281**, 20140018 (2014).
11. M. T. O'Mara *et al.*, *Curr. Biol.* **31**, 1311–1316.e4 (2021).
12. M. Teague O'Mara, M. Wikelski, B. Kranstauber, D. K. N. Dechmann, *Ecology* **100**, e02762 (2019).
13. R. Hutterer, T. Ivanova, C. Meyer-Cords, L. Rodrigues, *Bat Migrations in Europe: A Review of Banding Data and Literature* (BfN-Schriftenvertrieb im Landwirtschaftsverlag, 2005).
14. L. Printz, M. Tschapka, A. Vogeler, *J. Urban Econ.* **7**, juab033 (2021).
15. D. K. N. Dechmann *et al.*, *PLOS ONE* **9**, e114810 (2014).
16. R. P. Wilson *et al.*, *J. Anim. Ecol.* **89**, 161–172 (2020).
17. T. A. Wild *et al.*, *Anim. Biotelem.* **11**, 13 (2023).
18. E. Hurme, I. Lenzi, M. Wikelski, T. A. Wild, D. K. N. Dechmann, Data from: "Bats surf storm fronts during spring migration," Movebank Data Repository (2024); https://doi.org/10.5441/001/1.319.
19. A. Hedenström, *J. Mammal.* **90**, 1298–1309 (2009).
20. A. M. Bagg, *Auk* **65**, 147 (1948).
21. L. Qasem *et al.*, *PLOS ONE* **7**, e31187 (2012).
22. J. H. Jensen, J. C. Dyre, T. Hecksher, *PLOS ONE* **19**, e0303834 (2024).
23. P. A. Racey, *J. Reprod. Fertil.* **19**, 465–474 (1969).
24. W. F. Frick *et al.*, *Biol. Conserv.* **209**, 172–177 (2017).
25. E. Hurme, I. Lenzi, M. Wikelski, T. A. Wild, D. K. N. Dechmann, Code for: Bats surf storm fronts during spring migration, Zenodo (2024); https://doi.org/10.5281/zenodo.13910030.

### ACKNOWLEDGMENTS

We thank M. and F. Heeb and the Fledermausschutz Thurgau for assistance with monitoring and catching bats. Additionally, we thank ARA Frauenfeld for providing access to their facilities. This work was also made possible by many volunteers who helped in bat capture and processing, including N. Hwang, C. Monteza, J. Kohles, L. Keicher, C. Baldoni, E. Ye, A. Prylutska, N. Toshkova, A. Goldshtein, A. Boonman, N. Richter, and many more. We also thank T. O'Mara and J. Koblitz for helpful comments on the manuscript. **Funding:** This study was funded by the Max Planck Institute of Animal Behavior, the ICARUS initiative, Germany's Excellence Strategy-EXC 2117-422037984, the Gordon and Betty Moore Foundation through grant GBMF10539 to M.W., and the Academy for the Protection of Zoo Animals and Wildlife e.V. **Author contributions:** Conceptualization: E.H., T.A.W., M.W., D.K.N.D. Methodology: E.H., T.A.W., M.W., D.K.N.D. Investigation: E.H., I.L., M.W., D.K.N.D. Visualization: E.H. Funding acquisition: M.W., D.K.N.D. Project administration: E.H., D.K.N.D. Supervision: T.A.W., M.W., D.K.N.D. Writing – original draft: E.H., D.K.N.D. Writing – review and editing: E.H., T.A.W., I.L., M.W., D.K.N.D. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available through the Movebank Data Repository (*18*). Code is available at Zenodo (*25*). **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.ade7441
Materials and Methods
Supplementary Text
Figs. S1 to S5
Tables S1 to S10
References (*26*–*39*)
MDAR Reproducibility Checklist

Submitted 26 February 2024; resubmitted 21 July 2024
Accepted 15 November 2024
10.1126/science.ade7441

---

ORGANIC CHEMISTRY

# Sulfenylnitrene-mediated nitrogen-atom insertion for late-stage skeletal editing of *N*-heterocycles

**Bidhan Ghosh†, Prakash Kafle†, Rishav Mukherjee†, Randall Welles, Deacon Herndon, Kenneth M. Nicholas, Yihan Shao, Indrajeet Sharma***

**Given the prevalence of nitrogen-containing heterocycles in commercial drugs, selectively incorporating a single nitrogen atom is a promising scaffold hopping approach to enhance chemical diversity in drug discovery libraries. We harness the distinct reactivity of sulfenylnitrenes, which insert a single nitrogen atom to transform readily available pyrroles, indoles, and imidazoles into synthetically challenging pyrimidines, quinazolines, and triazines, respectively. Our additive-free method for skeletal editing employs easily accessible, benchtop-stable sulfenylnitrene precursors over a broad temperature range (−30 to 150°C). This approach is compatible with diverse functional groups, including oxidation-sensitive functionalities such as phenols and thioethers, and has been applied to various natural products, amino acids, and pharmaceuticals. Furthermore, we have conducted mechanistic studies and explored regioselectivity outcomes through density functional theory calculations.**

Scaffold hopping, first introduced by Schneider and colleagues (*1*), began as a virtual strategy in drug discovery to identify novel pharmacophores by making simple structural modifications that alter a molecule's biological and pharmacological properties (*2*, *3*). In recent years, several scaffold hopping methods have been developed for the direct editing of drug molecules. (*4*). Among them, skeletal editing has taken center stage as it modifies the molecular frameworks by adding or removing individual atoms (*5*). In the late-stage functionalization of scaffolds, notable progress has been achieved by editing carbon, nitrogen, or oxygen atoms within a skeleton (*5*–*12*). This process enhances the chemical diversity of existing libraries, allowing access to uncharted regions of chemical space in drug discovery (*13*).

Nitrogen atoms play a pivotal role in medicinal chemistry and drug development (*14*, *15*). For example, pyrroles, indoles, and imidazoles are prevalent nitrogen-containing heterocycles, and direct incorporation of a single nitrogen atom into these structures offers a straightforward route to synthetically challenging heterocycles (quinazoline, pyrimidine, or triazine) without altering the rest of the molecule. Therefore, incorporating a single nitrogen atom could enhance late-stage diversification by adding key pharmacophores to drug discovery libraries (*16*).

The strategy of *N*-atom insertion into an aromatic compound goes back to 1964 when ammonium chloride was used as the source for a single nitrogen atom in the presence of oxidizing agents (*17*) (Fig. 1A). Despite advances in direct *N*-atom insertion (*18*–*28*), most methods rely on nitrene intermediates that require harsh reaction conditions to generate, involving heavy metals (*25*) or excessive use of potent oxidizing (*24*) or pyrophoric reagents (*28*), which constrains synthetic applicability. Additionally, some methods require the protection of free nitrogen in substrates such as indoles, necessitating a separate prefunctionalization step (*27*). Moreover, the direct insertion of nitrogen atoms into pyrroles, azaindoles, and imidazoles remains underexplored, with only a limited number of reports (*28*, *29*).

## Reaction development

Motivated by the potential of nitrenes bearing a leaving group to incorporate a single nitrogen atom, we aimed to uncover a mild protocol that does not depend on oxidizing agents. Our investigation led us to sulfenylnitrenes, which possess a thio-functionality known for its leaving group capabilities. Despite their discovery in 1967 (*30*), these nitrenes have not been extensively explored (*23*, *31*–*35*). Following a comprehensive literature survey, we identified sulfenylnitrene precursors (SNPs) **SNP-1**, **SNP-2**, and **SNP-3**, capable of generating sulfenylnitrenes over a broad temperature range (Fig. 1B). These thermal precursors have only been used in a few prior reports on aziridination (*32*, *33*, *36*, *37*).

Building upon the literature precedents, we hypothesized that the aziridine intermediate featuring an inbuilt thio-moiety as a leaving group would facilitate ring expansion and enable the selective incorporation of a single nitrogen atom. The S–N bond length in sulfenylnitrenes is approximately 1.51 Å, indicating the double bond character between sulfur and nitrogen (*32*). This implies that the metallomimetic properties of sulfur not only stabilize sulfenylnitrenes but also modulate their reactivity for a variety of synthetic transformations (Fig. 1B).

---

Department of Chemistry and Biochemistry, University of Oklahoma, 101 Stephenson Parkway, Norman, OK, USA.
***Corresponding author. Email: isharma@ou.edu**

†These authors contributed equally to this work.

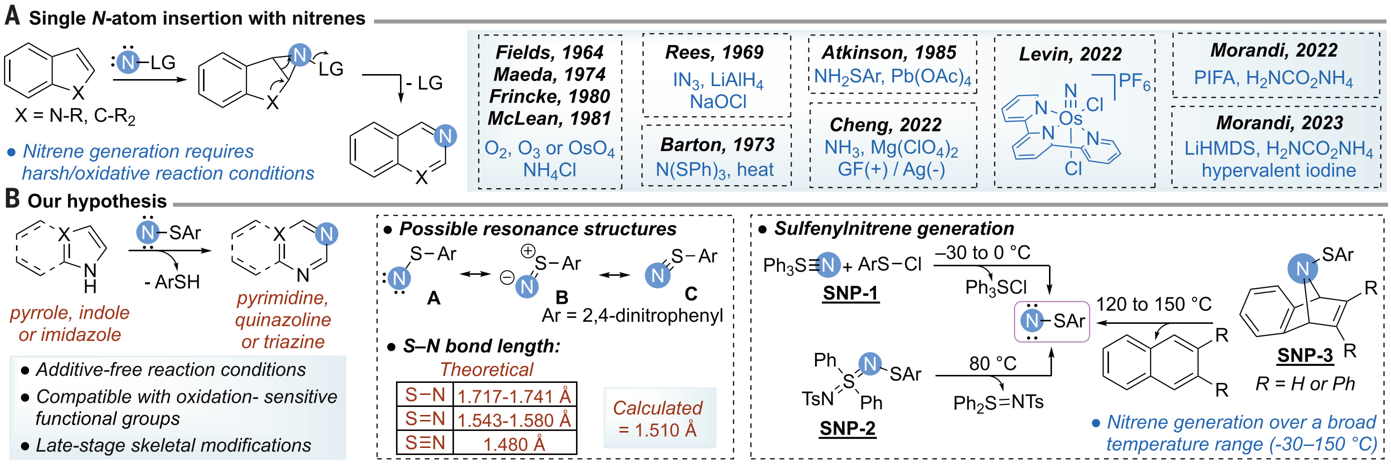
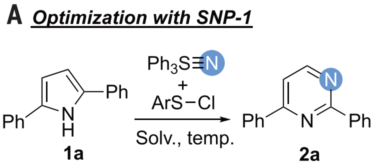
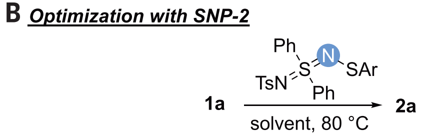
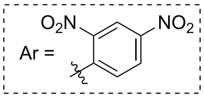


**Fig. 1. Background and design.** (**A**) Literature precedents for single *N*-atom insertion using nitrenes. (**B**) Our design using sulfenylnitrenes generated from SNPs to achieve tunable, protecting group-free single *N*-atom insertion over a broad temperature range (–30 to 150°C).

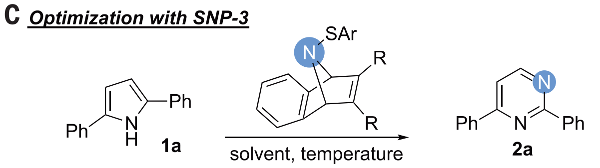


| Entry | Solv, Temp | Yield* |
|---|---|---|
| 1 | $CH_2Cl_2$, -30 °C | 15% |
| 2 | PhCl, -30 °C | 30% |
| **3** | **PhCl, 0 °C** | **50%** |
| 4 | PhCl, rt | 32% |





| Entry | Solvent | Yield* |
|---|---|---|
| 1 | acetonitrile | 77% |
| 2 | dichloroethane | 75% |
| 3 | toluene | 80% |
| **4** | **chlorobenzene** | **99%** |
| 5 | ethyl acetate | 73% |
| 6 | *t*-butanol | 48% |
| 7 | isopropanol | 40% |
| 8 | ethanol | 45% |
| 9.† | chlorobenzene | 68% |
| 10.‡ | chlorobenzene | 82% |



| Entry | SNP | Solvent | Temperature | Yield* |
|---|---|---|---|---|
| 1 | SNP-3a | chlorobenzene | 100 °C | 70% |
| 2 | **SNP-3b** | **chlorobenzene** | **120 °C** | **99%** |
| **3** | **SNP-3c** | **chlorobenzene** | **150 °C** | **99%** |
| 4 | SNP-3b | toluene | 120 °C | 81% |
| 5 | SNP-3b | xylene | 120 °C | 92% |
| 6 | SNP-3b | *n*-butanol | 120 °C | 50% |
| 7 | SNP-3b | *n*-butyl acetate | 120 °C | 50% |

**SNP-3a:** R = H, Ar = 4-nitrophenyl

**SNP-3b:** R = H, Ar = 2,4-dinitrophenyl

**SNP-3c:** R = Ph, Ar = 2,4-dinitrophenyl

**Fig. 2. Optimization of the *N*-insertion reaction.** (**A**) to (**C**) Reactions were carried out on a 0.03 to 0.1 mmol scale in a sealed vial with 2.0 equiv. of SNPs. *Yield was determined by $^1$H-NMR using 1,3,5-trimethoxybenzene as an internal standard. †1 equiv. of SNP-2 was used. ‡1.5 equiv. of SNP-2 was used.

To insert a single nitrogen atom into pyrroles, indoles, and imidazoles, we envisioned using benchtop-stable precursors of sulfenylnitrenes, **SNP-1**, **SNP-2**, and **SNP-3**. These precursors allow the generation of nitrenes under additive-free conditions over a broad temperature range (–30 to 150°C) as described in Fig. 1B (see fig. S8 in the SM for differential scanning calorimetry analyses).

Our optimization studies commenced with the use of symmetric 2,5-diphenyl pyrrole **1a** as the model substrate to overcome the regioselectivity challenge. We employed two equivalents of **SNP-1**, which was generated in situ by the reaction of triphenyl sulfanenitrile with commercially available 2,4-dinitrobenzenesulfenyl chloride at low-temperature using dichloromethane as a solvent (Fig. 2A), following literature precedent for the aziridination (*37*). The reaction produced the corresponding pyrimidine **2a** at –30°C, albeit in a modest yield (15%, entry 1). To enhance the yield, we switched to chlorobenzene as the solvent, inspired by its successful application in aziridination by Atkinson and coworkers (*35*), and the yield improved to 30%. Further raising the reaction temperature to 0°C resulted in an increased yield of 50%; however, temperatures above this led to increased byproduct formation. We next examined the reaction with **SNP-2**. In this case, acetonitrile was initially chosen as the solvent, following the literature precedent for aziridination (*32*). Notably, the reaction resulted in the formation of the corresponding pyrimidine **2a** in 77% yield (Fig. 2B, entry 1). Changing the solvent to dichloromethane or toluene did not substantially improve the yield (entries 2 and 3). However, using chlorobenzene as the solvent led to a nearly quantitative yield of pyrimidine **2a** (99%) (entry 4).

We also tested industrially preferred solvents, including esters and alcohols, and found that the reaction proceeded in both with moderate yields (entries 5 to 8). Finally, we screened 1 and 1.5 equivalents of **SNP-2** precursors, but this resulted in slightly lower yields and left some pyrroles unreacted (entries 9 and 10). Moving forward, we investigated the reactivity of various **SNP-3** precursors in the *N*-insertion reaction with pyrrole **1a** (Fig. 2C). Three distinct nitrene precursors were synthesized and treated with pyrrole **1a**. **SNP-3a** demonstrated a moderate yield (70%) at 100°C (*33*) when chlorobenzene was used as the reaction solvent (entry 1). By contrast, **SNP-3b** and **SNP-3c** exhibited better reactivity with pyrrole **1a**, yielding the desired pyrimidine **2a** in nearly quantitative yield, with chlorobenzene as the solvent (entries 2 and 3). The decomposition temperatures of **SNP-3b** and **SNP-3c** were determined to be

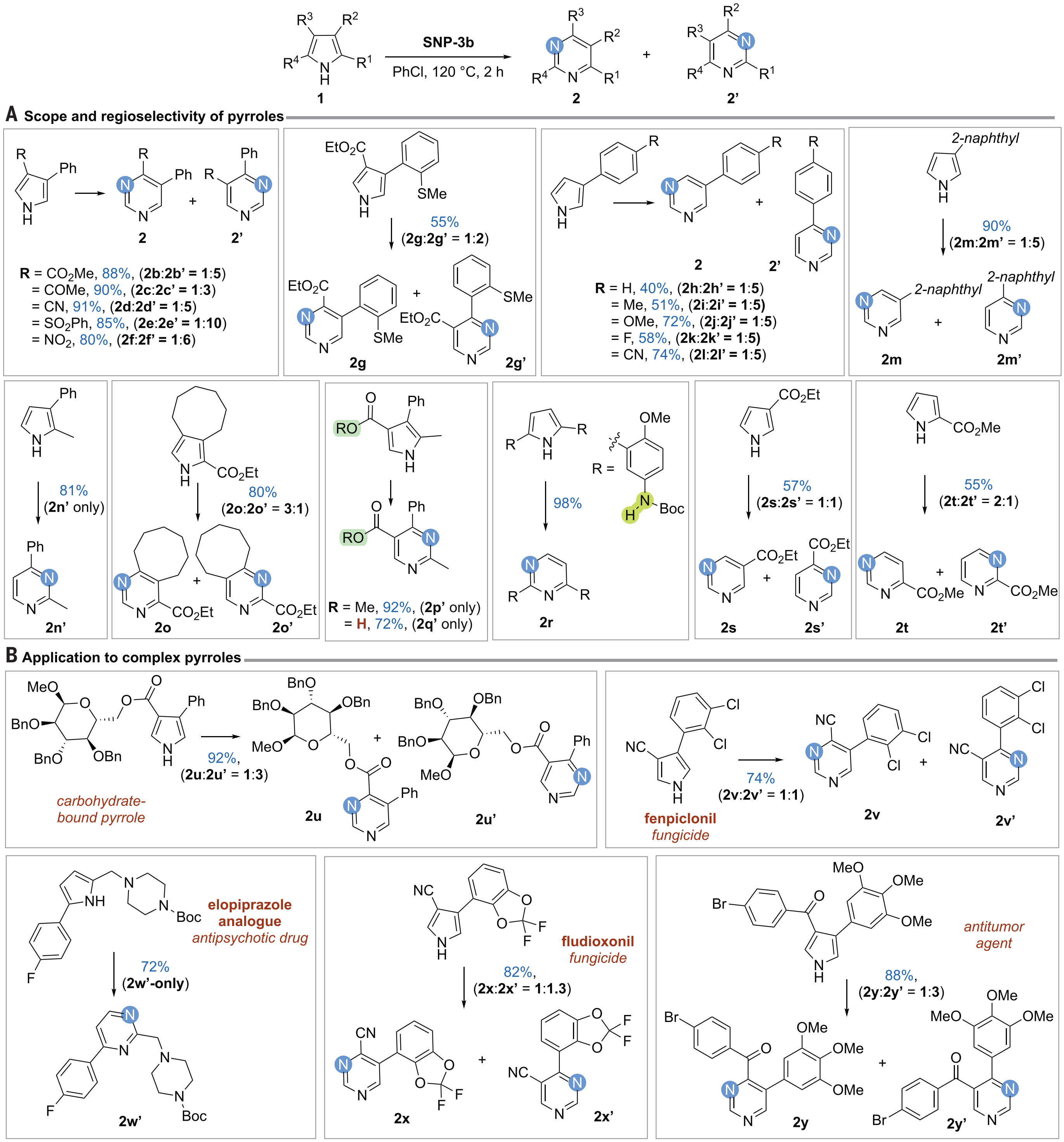


**Fig. 3. Scope and regioselectivity.** (**A**) Reaction with pyrroles. Conditions: 1 (1 equiv.), SNP-3b (2 equiv.), chlorobenzene (0.1 M), 120°C, two hours. Reactions were carried out on a 0.03 to 0.1 mmol scale and yields reported are isolated yields, except for the products having regioisomeric ratios, which were determined by $^1$H-NMR using 1,3,5-trimethoxybenzene as an internal standard. (**B**) Application to complex pyrroles.

120°C and 150°C, respectively. Other aromatic solvents, such as toluene and xylene, provided slightly lower yields (entries 4 and 5). However, the industrially preferred solvents (esters and alcohols) provided moderate yields of the desired pyrimidine **2a** (entries 6 and 7).

We selected **SNP-3b** as the optimal nitrene precursor due to the generation of sublimable naphthalene as a byproduct, easing the purification process at scale. Additionally, **SNP-3b** can be easily synthesized at a large scale (~10 to 20 g) in three simple steps (see SM for additional details).

## Scope and application of *N*-insertion into pyrroles

With these optimized conditions, we next investigated the scope and regioselectivity of pyrroles (Fig. 3A). The *N*-insertion preferentially occurs adjacent to electron-donating over

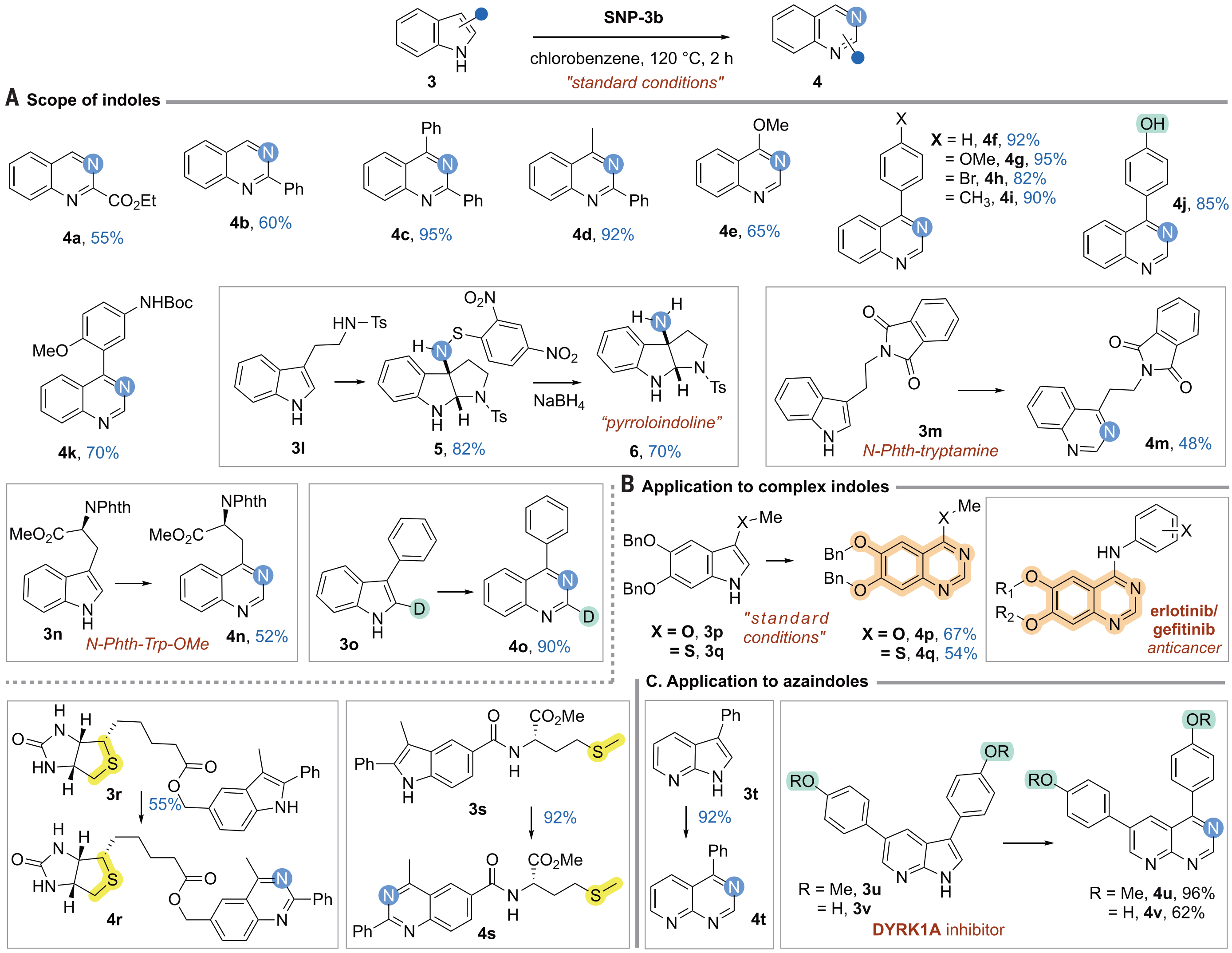


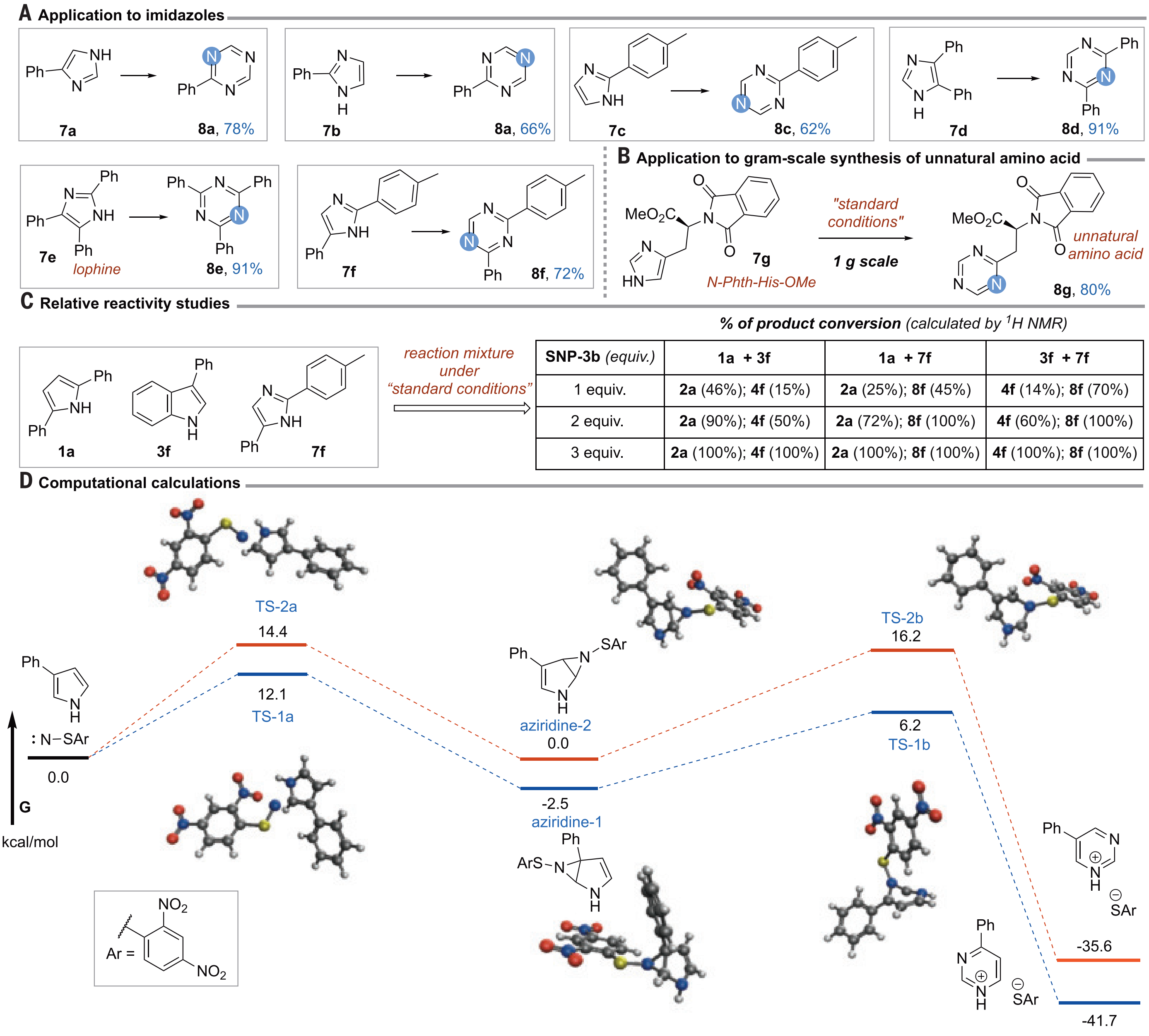
**Fig. 4. Application to indoles and azaindoles.** (**A**) Scope of indoles. Conditions: 3 (1 equiv.), SNP-3b (2 equiv.), chlorobenzene (0.1 M), 120°C, two hours. Reactions were carried out on a 0.03 to 0.1 mmol scale and yields reported are isolated yields. (**B**) Application to complex indoles. (**C**) Application to azaindoles.

electron-withdrawing substituents. For instance, a pyrrole featuring an aryl group at the 3-position and various electron-withdrawing substituents at the 4-position consistently produced the corresponding pyrimidine in good to excellent yields favoring *N*-insertion next to the aryl side (**2b** to **2f**). We observed that a highly oxidation–sensitive thioether functional group underwent the expansion reaction smoothly with a satisfactory yield (**2g**). Next, various 3-substituted pyrroles were subjected to the reaction conditions, and the *N*-insertion predominantly occurred from the substituted side of pyrrole. We observed that phenyl rings bearing electron-donating and electron-withdrawing groups at the 3-position of pyrrole did not affect the regioselectivity outcome (**2h** to **2l**). As anticipated, 3-naphthyl-pyrrole produced the corresponding pyrimidine in high yield (90%) with 1:5 regioselectivity (**2m**). We observed that *di-* and *tri-*substituted pyrroles predominantly yielded a single regio-isomer, favoring the more substituted position (**2n** to **2r**). A pyrrole derivative featuring an 8-membered carbocycle furnished the corresponding pyrimidine in high yield (**2o**). We observed that *tri*-substituted pyrroles bearing a methyl ester and a carboxylic acid functionality were compatible (**2p** and **2q**). Furthermore, pyrrole with a Boc-protected aromatic amine underwent the ring expansion reaction with a nearly quantitative yield (**2r**), highlighting the robustness of this method. Pyrroles with electron-withdrawing esters provided lower regioselectivity, favoring the electron-rich side (**2s** to **2t**).

To further expand the utility of this methodology, we applied it to complex bioactive pyrroles (Fig. 3B). We found that a carbohydrate-bound pyrrole was compatible preserving the stereochemistry at the sensitive anomeric acetal functionality (**2u**). The reaction also proved effective with fenpiclonil, an agricultural phenylpyrrole fungicide (*38*), yielding the expansion product in a good yield (**2v**). Additionally, the antipsychotic drug elopiprazole (*39*) demonstrated compatibility with the reaction, providing a yield of 72%, with only one regioisomer obtained (**2w**). Furthermore, fludioxonil, a fungicide used for seed treatment (*40*), was successfully transformed to the corresponding pyrimidine (**2x**). Finally, a trimethoxy-phenyl substituted pyrrole, recognized as an antitumor agent (*41*), successfully expanded into the corresponding pyrimidine (**2y**) with a high yield of 88%.

### Application to indole and azaindole scaffolds

Next, we harnessed the potential of sulfenylnitrenes to add a single *N*-atom to indoles, an important motif in numerous bioactive molecules (*42*) (Fig. 4A). Contrary to the literature reports, in which indole *N*-protection was necessary for the *N*-insertion, our methodology did not require any protection (*27*). This methodology



**% of product conversion** (calculated by $^1$H NMR)

| **SNP-3b** (equiv.) | **1a** + **3f** | **1a** + **7f** | **3f** + **7f** |
|---|---|---|---|
| 1 equiv. | **2a** (46%); **4f** (15%) | **2a** (25%); **8f** (45%) | **4f** (14%); **8f** (70%) |
| 2 equiv. | **2a** (90%); **4f** (50%) | **2a** (72%); **8f** (100%) | **4f** (60%); **8f** (100%) |
| 3 equiv. | **2a** (100%); **4f** (100%) | **2a** (100%); **8f** (100%) | **4f** (100%); **8f** (100%) |

**Fig. 5. Application to imidazoles and relative reactivity studies.** (**A**) Applications to imidazoles. Conditions: 7 (1 equiv.), SNP-3b (2 equiv.), chlorobenzene (0.1 M), 120°C, two hours. Reactions were carried out on a 0.03 to 0.1 mmol scale and yields reported are isolated yields. (**B**) Gram-scale synthesis of unnatural amino acid. (**C**) Relative reactivity studies of pyrrole, indole, and imidazole. (**D**) Computational calculations. Free energy profile for reaction of 3-Ph pyrrole with sulfenylnitrene **SNP-3b**. Energies are in kcal/mol; atomic colors are as follows: C, gray; O, red; N, blue; S, yellow; H, light gray.

accommodated various indoles, bearing both electron-donating and electron-withdrawing groups, and provided the corresponding quinazolines in good to high yields (**4a** to **4i**). The method also exhibited compatibility with the oxidation-sensitive functionality phenol (**4j**), which was incompatible with the existing methods as they require potent oxidizing agents. An indole containing a Boc-protected amine was compatible (**4k**). The reaction with tosyl-protected tryptamine (**3l**) provided the corresponding pyrroloindoline **5** as a single diastereomer in high yield. Mechanistically, the opening of the aziridine intermediate was favored over ring expansion, occurring from the same side of the aliphatic chain at the 3-position of tryptamine, resulting in exclusive *cis*-selectivity (see SM for additional details). This suggests that sulfenylnitrenes can also be used to access bioactive pyrroloindolines, a diverse family of structurally complex polyindoline alkaloids derived from various natural sources including amphibians, plants, and marine algae (*43*). Additionally, the N–S bond of the thioether functionality was easily cleaved through reduction with sodium borohydride. To avoid the cyclization product pyrroloindoline, we used phthalimide-protected tryptamine and tryptophan and transformed them into their respective quinazoline products (**4m** and **4n**), thereby providing access to previously underexplored unnatural amino acids and their derivatives (*44*). We also attempted the reaction with 2-deuterated indole to understand the mechanism further. As expected, no loss of deuterium was observed in the product (**4o**), supporting the mechanistic hypothesis described in the computational simulations (see Fig. 5C and fig. S9).

We further demonstrated the potential of this method for synthesizing readily accessible starting materials for medicinally relevant compounds, such as the anticancer agents erlotinib and gefitinib (*45*), which could aid in developing a robust library of drugs leads. To access their analogs, a substituted indole bearing a thioether functionality **3q** was transformed into quinazoline **4q**, which could be readily converted into erlotinib or gefitinib analogs as outlined in the literature (*46*) (Fig. 4B). Biotinylated indoles serve as bifunctional probes for detecting indole-binding proteins, which are integral to numerous biological processes (*47*). Additionally, 23% of newly FDA-approved small-molecule drugs from 2013 to 2023 feature sulfur moieties (*48*). To demonstrate the compatibility with oxidation-sensitive sulfur-containing complex substrates, a biotin-containing indole **3r** was successfully transformed into the corresponding quinazoline **4r** with good yield. The reaction also showed effectiveness with methionine-derivatized indole **3s**. These results showcase the compatibility of our nitrene reaction conditions with various functionalities including urea, amide, and thioether.

Furthermore, azaindoles are crucial core structures in pharmaceuticals and natural products, widely used in various areas of medicinal chemistry (*49*). However, existing *N*-insertion methods have proven inadequate for expanding azaindoles. Our methodology successfully expanded azaindoles to the corresponding 1,3,8-triazanaphthalenes (Fig. 4C; **4t** and **4u**) with high efficiency. We observed that a phenol-containing azaindole **3v**, a DYRK1A inhibitor (*50*), underwent a smooth *N*-insertion reaction, furnishing the desired product in high yield.

### Application to imidazoles, relative reactivity studies, and computational calculations

Following the successful implementation of *N*-insertion into pyrroles and indoles, we also extended its application to imidazoles (Fig. 5A). We observed that various aryl-substituted imidazoles smoothly converted into the corresponding 1,3,5-triazenes, a privileged motif in drug discovery (**8a** to **8f**). Lophine, a chemiluminescent molecule known for its prolonged luminescence, readily underwent expansion in excellent yield (**8e**). Additionally, the phthalimide-protected L-histidine methyl ester (**7g**) was successfully transformed into the corresponding unnatural triazine amino acid (**8g**) at the gram scale (Fig. 5B).

We also investigated the relative reactivity of pyrrole, indole, and imidazole with sulfenylnitrene precursor **SNP-3b**. Combinations of equimolar amounts of pyrrole (**1a**), indole (**3f**), and imidazole (**7f**) were subjected to the standard reaction conditions with 1.0, 2.0, and 3.0 equiv. of **SNP-3b** (Fig. 5C). Rate of *N*-insertion with 1.0 and 2.0 equiv. **SNP-3b** followed the order imidazole > pyrrole > indole (see figs. S21 to S23 for details). Notably, 3 equiv. of **SNP-3b** fully converted all combinations to the desired products.

Finally, we sought to elucidate the pathway for *N*-insertion to pyrroles by conducting density functional theory (DFT) calculations (B3LYP//6-31G(d)//D3-BJ//SMD-PhCl), resulting in the energy profile shown in Fig. 5D. The reaction begins with the rate-limiting release of 2,4-dinitrophenylsulfenylnitrene, likely a ground state singlet, from the precursor **SNP-3b** ($\Delta G^{*}$ = 31 kcal/mol, not shown). The sulfenylnitrene can react with a substituted pyrrole through asymmetric dipolar transition states **TS-1a** or **TS-2a** to form **aziridine-1** or **aziridine-2**. In this asynchronous concerted process, *C2/C5*-bond formation precedes *C3/C4*-bond formation. The insertion regioselectivity, which favors *N*-attack at *C2* over *C5*, can be explained by the more effective delocalization of the developing positive charge in **TS-1a** by the phenyl group at the *C3* position. The reverse regioselectivity observed for substrates **2s** and **2t** can also be explained by our DFT model. Here, the nitrene will preferentially add to the more electron rich side of the pyrrole, favoring the dipolar structure that avoids developing a positive charge adjacent to the electron-withdrawing ester group (see figs. S24 to S31 for details).

These SNPs can be synthesized at a large scale, highlighting their potential for commercial availability. Additive-free nitrene generation, an operationally simple protocol, and good functional group compatibility highlight the advantages of sulfenylnitrenes in creating distinct chemical entities to explore the uncharted chemical space in drug discovery.

## REFERENCES AND NOTES

1. G. Schneider, W. Neidhart, T. Giller, G. Schmid, *Angew. Chem. Int. Ed.* **38**, 2894–2896 (1999).
2. Y. Hu, D. Stumpfe, J. Bajorath, *J. Med. Chem.* **59**, 4062–4076 (2016).
3. P. R. Lazzara, T. W. Moore, *RSC Med. Chem.* **11**, 18–29 (2019).
4. K. R. Campos *et al.*, *Science* **363**, eaat0805 (2019).
5. J. Jurczyk *et al.*, *Nat. Synth.* **1**, 352–364 (2022).
6. S. H. Kennedy, B. D. Dherange, K. J. Berger, M. D. Levin, *Nature* **593**, 223–227 (2021).
7. G. L. Bartholomew, F. Carpaneto, R. Sarpong, *J. Am. Chem. Soc.* **144**, 22309–22315 (2022).
8. B. W. Joynson, L. T. Ball, *Helv. Chim. Acta* **106**, e202200182 (2023).
9. Q.-H. Shi *et al.*, *Org. Chem. Front.* **10**, 3307–3312 (2023).
10. M. Peplow, *Nature* **618**, 21–24 (2023).
11. P. Zhang, L. Hua, T. Takahashi, S. Jin, Q. Wang, *Synthesis* **56**, 55–70 (2024).
12. Q. Cheng *et al.*, *Nat. Chem.* **16**, 741–748 (2024).
13. I. Sharma, D. S. Tan, *Nat. Chem.* **5**, 157–158 (2013).
14. N. Kerru, L. Gummidi, S. Maddila, K. K. Gangu, S. B. Jonnalagadda, *Molecules* **25**, 1909 (2020).
15. M. M. Heravi, V. Zadsirjan, *RSC Adv.* **10**, 44247–44311 (2020).
16. Y. Hu, D. Stumpfe, J. Bajorath, *J. Med. Chem.* **60**, 1238–1246 (2017).
17. M. I. Fremery, E. K. Fields, *J. Org. Chem.* **29**, 2240–2243 (1964).
18. D. C. Horwell, C. W. Rees, *J. Chem. Soc. D* **23**, 1428a (1969).
19. D. H. Barton, I. A. Blair, P. D. Magnus, R. K. Norris, *J. Chem. Soc., Perkin Trans. 1*1037–1040 (1973).
20. K. Maeda, T. Mishima, T. Hayashi, *Bull. Chem. Soc. Jpn.* **47**, 334–338 (1974).
21. R. B. Miller, J. M. Frincke, *J. Org. Chem.* **45**, 5312–5315 (1980).
22. D. S. Dime, S. McLean, *J. Org. Chem.* **46**, 4999–5000 (1981).
23. R. S. Atkin son, B. D. Judkins, D. R. Russell, L. J. Sherry, *J. Chem. Soc., Perkin Trans. 1*1967–1969 (1985).
24. P. R. Kumar, *Heterocycles* **26**, 1257–1262 (1987).
25. P. Q. Kelly, A. S. Filatov, M. D. Levin, *Angew. Chem. Int. Ed.* **61**, e202213041 (2022).
26. S. Liu, X. Cheng, *Nat. Commun.* **13**, 425 (2022).
27. J. C. Reisenbauer, O. Green, A. Franchino, P. Finkelstein, B. Morandi, *Science* **377**, 1104–1109 (2022).
28. J. C. R *et al.*, *Org. Lett.* **25**, 8419–8423 (2023).
29. Y. Shinkai, "Hetarenes and Related Ring Systems" in *Science of Synthesis*, vol. 17 (Thieme Publishing Group, 2004).
30. F. Anet, R. D. Trepka, D. J. Cram, *J. Am. Chem. Soc.* **89**, 357–362 (1967).
31. T. Q. Davies *et al.*, *J. Am. Chem. Soc.* **142**, 15445–15453 (2020).
32. T. Yoshimura, T. Fujie, T. Fujii, *Tetrahedron Lett.* **48**, 427–430 (2007).
33. R. S. Atkinson, B. D. Judkins, *J. Chem. Soc., Perkin Trans. 1* 2615–2619 (1981).
34. X. Tian *et al.*, *Angew. Chem. Int. Ed.* **58**, 3589–3593 (2019).
35. R. S. Atkinson, M. Lee, J. R. Malpass, *J. Chem. Soc. Chem. Commun.* **14**, 919–920 (1984).
36. A. Lork *et al.*, *J. Chem. Soc., Chem. Commun.* **12**, 898–899 (1992).
37. T. Fujii, T. Kousaka, T. Yoshimura, *Tetrahedron Lett.* **43**, 5841–5843 (2002).
38. A. B. Jespers, L. C. Davidse, M. A. Dewaard, *Pestic. Biochem. Physiol.* **45**, 116–129 (1993).
39. I. van Wijngaarden, C. G. Kruse, R. van Hes, J. A. van der Heyden, M. T. Tulp, *J. Med. Chem.* **30**, 2099–2104 (1987).
40. T. T. Brandhorst, I. R. L. Kean, S. M. Lawry, D. L. Wiesner, B. S. Klein, *Sci. Rep.* **9**, 5047 (2019).
41. X. P. Zhan *et al.*, *Chem. Biodivers.* **14**, e1600219 (2017).
42. N. K. Kaushik *et al.*, *Molecules* **18**, 6620–6662 (2013).
43. C. Sun, W. Tian, Z. Lin, X. Qu, *Nat. Prod. Rep.* **39**, 1721–1765 (2022).
44. T. Narancic, S. A. Almahboub, K. E. O'Connor, *World J. Microbiol. Biotechnol.* **35**, 67 (2019).
45. M. A. Bareschino *et al.*, *Ann. Oncol.* **18**, vi35–vi41 (2007).
46. V. Chandregowda, G. V. Rao, G. C. Reddy, *Synth. Commun.* **37**, 3409–3415 (2007).
47. E. Dolusić, M. Kowalczyk, V. Magnus, G. Sandberg, J. Normanly, *Bioconjug. Chem.* **12**, 152–162 (2001).
48. C. M. Marshall, J. G. Federice, C. N. Bell, P. B. Cox, J. T. Njardarson, *J. Med. Chem.* **67**, 11622–11655 (2024).
49. D. R. Motati, R. Amaradhi, T. Ganesh, *Bioorg. Med. Chem.* **28**, 115830 (2020).
50. S. Gourdain *et al.*, *J. Med. Chem.* **56**, 9569–9585 (2013).

## ACKNOWLEDGMENTS

We thank N.G. Akhmedov and S. Foster from Research Support Services at the University of Oklahoma for NMR and mass spectral analyses, respectively. We gratefully acknowledge B.P. Grady and H. Aboukeila from the Department of Chemical, Biological, and Materials Engineering at the University of Oklahoma for their assistance with the DSC analyses. We also acknowledge support from the OU Supercomputing Center for Education & Research (OSCER). **Funding:** This work was supported by NSF CHE-1753187 (to I.S.). **Author contributions:** I.S. conceptualized the idea and supervised the research. B.G., P.K., R.M., and R.W designed and conducted synthetic experiments, including purification and characterization. D.H., K.M.N, and Y.S. conducted computational studies. B.G. and I.S. prepared the manuscript with input from all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the supplementary materials.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adp0974
Materials and Methods
Supplementary Text
Figs. S1 to S31
Tables S1 to S3
References (*51*–*103*)

Submitted 13 March 2024; resubmitted 26 September 2024
Accepted 19 November 2024
10.1126/science.adp0974

### Degassing of aqueous solutions

Biotech Fluidics has introduced the DEGASi® life science degassing chamber that provides efficient degasification of aqueous liquids at flow rates up to 20 ml/min. Toxic poly- and perfluoroalkyl substances (PFAS) are known to be highly persistent in the environment and linked to health problems in humans, including higher cholesterol, lower birth weights, and kidney cancer. Traditionally most degassing systems have used fluorinated polymer membranes to ensure wide chemical compatibility. However, reflecting the growing demand from environmental, food, and healthcare labs to avoid PFAS, the new life science degassing chamber has been designed with a fluid path lined by a silicone membrane thereby eliminating the chance of contamination by these toxic substances. The DEGASi® life science degassing chamber is designed to improve analytical instrument precision and assay quality by removing dissolved gases from aqueous fluids before they outgas and form bubbles that impact precision dispensing or interfering with detection. The core functional element of the life science degassing chamber is a fluid path lined by a highly permeable silicone membrane. Fluid flows through the membrane pathway while the space surrounding the membrane is evacuated. A PID controlled vacuum pump holds a precise vacuum level driving the active removal of dissolved gases across the membrane. As a result, these degassing chambers provide a valuable tool for actively removing dissolved gases from aqueous fluid streams before they form bubbles, ensuring maximum analytical uptime and performance.

**Biotech Fluidics**
For info: +46-300-56-9180
www.biotechfluidics.com

### Animal Multiplex Assays

Rapid, accurate detection of six gastrointestinal disease-causing pathogens in cats and dogs. Primerdesign (part of the Novacyt Group), a company focussed on the design, manufacture, validation and supply of real-time PCR kits and reagents, announced that it is launching two new assays for gastrointestinal (GI) disease in dogs and cats. The genesig®PLEX kits for Canine and Feline Gastrointestinal Diseases are multiplexed qualitative PCR (qPCR) assays that detect six disease-causing pathogens in the digestive system of cats and dogs in a two-tube assay, delivering a streamlined and efficient testing solution for centralized veterinary laboratories. Each target has been carefully selected to detect a key pathogen that can cause potentially serious or life-threatening GI illnesses in cats and dogs. Testing for six pathogens using one sample will help to reduce hands-on time for laboratory technicians as well as the use of consumables. The kits provide results in less than three hours, enabling quicker, more comprehensive results to the veterinarian and ensuring faster treatment for the family pet. Adding to the 200+ veterinary assays in Primerdesign's existing portfolio, these tests demonstrate the company's continued commitment to identifying emerging diagnostic needs and providing best-in-class solutions for our customers.

**Primerdesign**
For info: +44 (0) 02380 748 830
www.primerdesign.co.uk

### DrySyn Vortex system

DrySyn Vortex Parallel Overhead Stirring System is now available with active temperature control. Responding to customer demand, the new DrySyn Vortex Temperature Controlled System is engineered to provide active heating or cooling in parallel when attached to a suitable heater/cooler, thus maximizing reaction conditions. The DrySyn Vortex is ideal for chemists who require over-head stirring capabilities where viscous or heterogeneous systems are used or where the use of magnetic stirring bars may damage materials due to their grinding effect. With three separate parallel reaction positions for up to 500 mL round bottom flasks all within a compact footprint, its space-saving design allows for reaction screening or performing multiple reactions. The incorporation of a circulator in the new DrySyn Vortex Temperature Controlled System allows active cooling or heating to maintain a constant solution temperature, facilitating controlled ramping to reduce overshoot. For those handling exothermic/endothermic reactions, or those carrying out polymerization, this new addition to the Asynt range offers exciting possibilities. With precise temperature control from ambient to 150 °C as standard, or -30 °C to 150 °C by using optional DrySyn Snow-Storm MULTI inserts, the system is highly versatile and efficient. Compatible with all major brands of overhead stirrer, users can triple the workload attributed to each motor and work both more efficiently and sustainably. The DrySyn Vortex system works by utilizing a robust maintenance-free mechanical gearbox which the user can easily adjust to provide more powerful stirring via a reduction drive or achieve better performance when using lower power overhead stirrers. The user can also easily swap their preferred stirrer shaft/head to suit their requirements, making this a long-lasting solution for the lab.

**Asynt**
For info: +44 (0)1638 781 709
www.asynt.com

### HPLC Pump Qualification Pack

TESTA Analytical has launched the HPLC Pump Qualification Pack, an efficient, fast, and reliable tool for regulated laboratories looking to qualify the operation and performance of their HPLC systems. Flow rate is one of the most important parameters in any liquid chromatography system, it determines retention time or volume and therefore has a major influence on reproducibility. The TESTA Analytical LC flowmeter is a trusted device for continuous non-invasive measurement of flow rate in HPLC and UHPLC systems. Comprising a calibrated LC flowmeter, data validation tool and tablet-computer with pre-installed software, the pack automatically performs all necessary flow data calculations plus operational and performance qualification checks. Using the HPLC Pump Qualification Pack considerably reduces the time required for HPLC pump qualification, while improving traceability of results. By focusing on data traceability and simplifying the qualification process, this new all-in-one pack is an affordable solution for laboratories seeking to maintain compliance and optimize their HPLC systems quickly and effectively.

**Testa Analytical Solutions**
For info: +49-30-864-24076
www.testa-analytical.com

---

Electronically submit your new product description or product literature information! Go to www.science.org/about/new-products-section for more information.

**Food and Agriculture Organization of the United Nations**

**The Food and Agriculture Organization of the United Nations (FAO) is currently seeking outstanding candidates for the position of Chief Scientist at its headquarters in Rome, Italy.**

Application deadline: 2 FEBRUARY 2025

FAO is a specialized agency of the United Nations that leads international efforts to defeat hunger.

FAO strives for a world free of hunger and malnutrition where food and agriculture contribute to the wellbeing of all, especially the poorest, in an economically, socially and environmentally sustainable fashion. The Organization proudly works in over 140 countries worldwide and believes that everyone can play a part in ending hunger, and there are many parts to play.

**Questions? Contact employment@fao.org**

The Chief Scientist makes a key contribution to the vision of Leaving No One Behind through harnessing Science, Technology and Innovation (STI) to provide solutions that address all the dimensions of sustainability to promote a world free from poverty, hunger and malnutrition. The Chief Scientist is the driver to leverage and mainstream STI at FAO by delivering strategic directions, enhancing dialogues and forging transformative partnerships. The Chief Scientist enables the Organization to strengthen its position as a source of reliable scientific information and a neutral platform at the heart of important debates, and strengthens FAO's capacities to support countries in identifying, piloting, and scaling up technologies and innovations which are adapted to their needs and contexts.

**Work with us and help to make a difference now!** Find out more about this opportunity on the Organization's careers page: http://www.fao.org/employment/vacancies/en

*We welcome qualified women and diverse talent worldwide to join us. Persons with disabilities are equally encouraged to apply.*

---

## Let *Science* Careers help advance your career.

- Register for a free online account on **ScienceCareers.org**.
- Search hundreds of job postings.
- Sign up to receive job alerts that match your criteria.
- Upload your resume into our database to connect with employers.
- Watch one of our many webinars on different career topics such as job searching, networking, and more.
- Download our career booklets, including Career Basics, Careers Beyond the Bench, and Developing Your Skills.
- Complete a personalized career plan at "my IDP."
- Read relevant career advice articles from our library of thousands.

SCIENCECAREERS.ORG

Visit **ScienceCareers.org** today — all resources are free

**ScienceCareers**
FROM THE JOURNAL SCIENCE AAAS

**WORKING LIFE**

*By* **Catherine Croft**

# The classroom door

When the call came for volunteers to visit local schools for Brain Awareness Week, I jumped at the opportunity. I had spent the past 3 years with my head down, focused on my neuroscience Ph.D., and I liked the idea of getting out and sharing what I was learning. The other graduate students and I brought a real human brain, using it as a guide to talk about the frontal lobe, hippocampus, amygdala, and other brain regions. "Where is creativity located?" one student asked. I found it deeply satisfying to help them understand even just a little bit about why their grandparents might not remember things or why the brains of their autistic brothers or sisters might work differently. The next day, I went back to my normal lab routine. But the experience planted a seed, eventually leading me to a new career.

I entered grad school with the same hope as many other students: to become a professor. I imagined myself running a research lab while teaching undergrads, inspiring them to become passionate about discovery.

I wasn't required to teach, but I lobbied my department to make me a teaching assistant. I also volunteered here and there for outreach activities by judging science fairs, giving lab tours, and visiting classrooms. But mostly I focused squarely on research, which I was told should be the priority if I wanted an academic job.

"I have loved finding ways to break down complicated concepts into simple parts."

I was perfectly happy with that for many years, as I loved making discoveries and publishing papers about them. But after I became a postdoc I realized something was off. I felt lost in the fine details of the research and yearned to see the big picture impact.

Amid my struggle, flyers and emails popped up announcing that my institute was going to hold its first ever career symposium. My curiosity was piqued. Could I find a meaningful career outside academia?

The education panel captivated me. It was eye opening to hear scientists describe how they had transitioned to positions at museums or research institutes, where they shared scientific knowledge with children and other members of the public. It was my first inkling that science education could even be a career. Afterward, I asked the panelists for advice on how I could follow in their footsteps, although I admitted, "I don't know if I'm brave enough."

The panelists convinced me I could make a career out of my interest in outreach. But after my fellowship ended I faced a new set of obstacles. I quickly realized my new path wasn't as straightforward as the academic route. "How do I find these jobs?" I wondered. I also worried about having to face the condescending looks of former colleagues, who I feared would view me as a failed postdoc. But, close to 40 years old at that point, with two young children, I chose to be courageous and do what was best for me.

I ended up following a winding path. Initially, I wrote about scientific discoveries for the public and worked at a science enrichment center. Eventually I found my way into a job teaching biology to public high school students. The first year was difficult, because I had to learn classroom management and best teaching practices. But I soon navigated my way through the ins and outs of my new profession and became confident in my choice.

In the 10 years I've spent as a teacher, I have loved finding ways to break down complicated concepts into simple parts. My science training has come in handy when guiding students through labs and helping them realize there is no "correct" answer to be found. And I have enjoyed telling them about my research experiences, including how often experiments fail. It's gratifying to think I've done my best to guide their lives in new directions, however big or small.

I sometimes look back wistfully on my research career. But when I see the immediate impact that I am making on young people's lives, I know my new path is a good match for my skills and passions. Scientists come in many forms and can serve society in different ways. It took me a long time to realize that, but I am at peace with it now. ■

---

Catherine Croft is a co-teacher at the Nysmith School and a co-founder of Catlilli Games. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER