How to share your data—and why you should p. 322

Specialized retinal neurons encode light intensity pp. 335 & 376

Origins of Earth's volatile elements pp. 369 & 372

Science

$15
27 JANUARY 2023
science.org

AAAS



AMAZON LOST

Forest degradation and destruction

pp. 348 & 349

MICHELSON MEDICAL RESEARCH FOUNDATION

Human Immunome Project

THE 2022 MICHELSON PRIZES

The Michelson Medical Research Foundation and the Human Immunome Project congratulate the recipients of the 2022 Michelson Prizes: Next Generation Grants.

The Michelson Prizes are $150,000 research grants given annually to support early-career researchers who are working to advance human immunology, vaccine discovery, and immunotherapy research for major global diseases.

Dr. Noam Auslander, **Dr. Jenna Guthmiller**, **Romain Guyon** and **Dr. Brittany Hartwell** were selected by a distinguished committee of internationally recognized scientists and represent the next generation of innovators in human immunology and vaccine research.

The recipeints of the 2022 Michelson Prizes will be honored at a virtual award ceremony in collaboration with Keystone Symposia on March 24, 2023 at 8:00 AM PST.



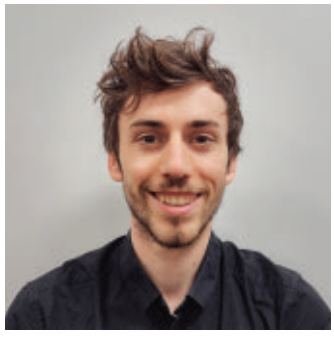
Noam Auslander, Ph.D.
Assistant Professor,
The Wistar Institute

"Artificial intelligence-based identification of microbes associated with immune responses in cancer and immune diseases."



Jenna Guthmiller, Ph.D.
Assistant Professor,
University of Colorado Anschutz School of Medicine

"Mapping preexisting mucosal B cell specificities engaged by potential universal influenza vaccines."



Romain Guyon
Ph.D. Student,
University of Oxford

"Demonstration of new single-dose vaccine technology towards a first-in-man clinical application."



Brittany Hartwell, Ph.D.
Assistant Professor,
University of Minnesota

"Engineering albumin-hitchhiking intranasal vaccines with enhanced transmucosal uptake to promote immunity."

COULD YOU BE THE NEXT MICHELSON PRIZE WINNER? APPLY FOR THE 2023 MICHELSON PRIZES.

Application portal opens April 3, 2023.

For more information visit:
www.humanimmunomeproject.org
www.michelsonmedicalresearch.org

CONTENTS



27 JANUARY 2023 • VOLUME 379 • ISSUE 6630

322

NEWS

IN BRIEF

314 News at a glance

IN DEPTH

316 Bird flu spread between mink is a 'warning bell'
Big outbreak at a Spanish farm reignites fears of an H5N1 influenza pandemic
By K. Kupferschmidt

317 Women at ocean science institute have half the lab space of men
Disparities exposed by "stunning" report from Scripps Institution of Oceanography are unlikely to be unique
By M. Wadman

318 Meteorite results bode well for exo-Earths
Water and light elements are likely available right where rocky planets take shape
By P. Voosen

320 Can California's floods help recharge depleted aquifers?
Plans to drown orchards and farm fields to boost groundwater supplies get off to a slow start *By D. Charles*

321 Fukushima wastewater release set to start soon
Japan, facing broad opposition, says dump will pose no risk to Pacific Ocean life or humans *By D. Normile*

FEATURES

322 Ready, set, share!
As funders roll out new requirements for making data freely available, researchers weigh costs and benefits
By J. Kaiser and J. Brainard

INSIGHTS

POLICY FORUM

326 Farm animal welfare: Beyond "natural" behavior
An animal-centered view guided by what animals value could improve welfare on farms
By M. S. Dawkins

PERSPECTIVES

329 Earth's persistent thermostat
The dependence of rock weathering on temperature helps to steer Earth's climate
By R. G. Hilton
RESEARCH ARTICLE p. 382

330 Decoding complex order in reticular frameworks
Periodicity and aperiodicity coexist in a nonchaotic, information-rich crystal structure *By O. M. Yaghi and Z. Rong*
RESEARCH ARTICLE p. 357

332 A vaccine sanctuary in the lymph node
B cell follicles in the lymph node protect vaccines to enhance immune responses
By A. Kalia and H. D. Hickman
RESEARCH ARTICLE p. 350

333 Infection infidelities drive innate immunity
Unsuccessful pathogenic activities may be key to activating host defenses to infection
By J. C. Kagan

335 Neurons share an intense load
Time-of-day–detecting cells in the eye give customized responses to light intensity
By P. R. Martin
RESEARCH ARTICLE p. 376

337 Dieter Oesterhelt (1940–2022)
Pioneer of membrane protein research
By P. Hegemann and H. Michel

BOOKS *ET AL.*

338 Science in the age of podcasts
A powerful 21st-century medium affords new opportunities for learning and creative thinking *By A. Gomez-Marin*



326

CREDITS: (ILLUSTRATION) DAVIDE BONAZZI/SALZMANART; (PHOTO) ANGEL GARCIA/BLOOMBERG/GETTY IMAGES

339 A population pessimist turns 90
Paul Ehrlich reflects on life, love, and the science and politics of the human predicament *By P. Gleick*

LETTERS

340 EU deforestation law overlooks emerging crops *By L. L. Powell* et al.

341 Hidden costs of Europe's deforestation policy *By G. W. Fernandes* et al.

341 Animal agriculture's potential financial risks *By D. Narain*

342 Technical Comment abstracts

RESEARCH

IN BRIEF

345 From *Science* and other journals

REVIEWS

Tropical forest

348 Human impacts outpace natural processes in the Amazon *J. S. Albert* et al.
REVIEW SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABO5003

349 The drivers and impacts of Amazon forest degradation *D. M. Lapola* et al.
REVIEW SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABP8622

RESEARCH ARTICLES

350 Vaccines
Low protease activity in B cell follicles promotes retention of intact antigens after immunization *A. Aung* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABN8934
PERSPECTIVE p. 332

351 Respiratory complexes
Structural basis of mammalian respiratory complex I inhibition by medicinal biguanides *H. R. Bridges* et al.

357 Crystalline ordering
Truchet-tile structure of a topologically aperiodic metal–organic framework *E. G. Meekel* et al.
PERSPECTIVE p. 330

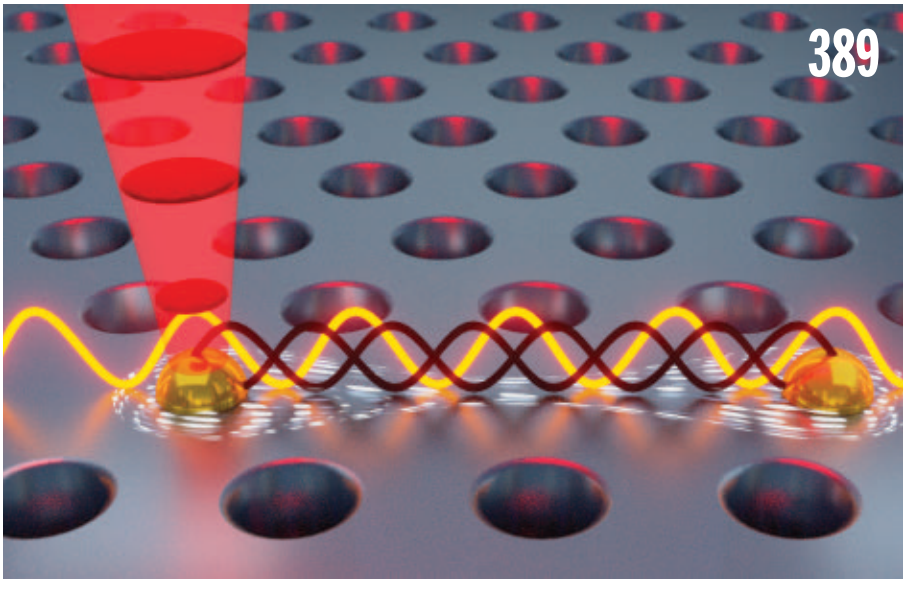
361 Biosynthesis
Complex scaffold remodeling in plant triterpene biosynthesis *R. De La Peña* et al.

Cosmochemistry

369 Nucleosynthetic isotope anomalies of zinc in meteorites constrain the origin of Earth's volatiles *R. Martins* et al.

372 Meteorites have inherited nucleosynthetic anomalies of potassium-40 produced in supernovae *N. X. Nie* et al.

376 Neuroscience
Encoding of environmental illumination by primate melanopsin neurons *A. Liu* et al.
PERSPECTIVE p. 335



Long-range photon-mediated coupling is demonstrated between two quantum dots in a photonic crystal.



382 Geochemistry
How temperature-dependent silicate weathering acts as Earth's geological thermostat *S. L. Brantley* et al.
PERSPECTIVE p. 329; PODCAST

389 Nanophotonics
Collective super- and subradiant dynamics between distant optical quantum emitters *A. Tiranov* et al.

393 Evolutionary ecology
Rapid plant trait evolution can alter coastal wetland resilience to sea level rise *M. L. Vahsen* et al.

399 Solar cells
Highly efficient p-i-n perovskite solar cells that endure temperature variations *G. Li* et al.

DEPARTMENTS

313 Editorial
ChatGPT is fun, but not an author *By H. H. Thorp*

406 Working Life
Finding time for my passion *By H. Merali*

ON THE COVER

The Amazon forest is changing rapidly as a result of human activities, including deforestation for agriculture, such as these soybean fields in Belterra, Pará, Brazil. Remaining areas of forest are experiencing an increased incidence of fires, drought, and the effects of neighboring land uses. These changes threaten local biodiversity and communities and alter the global climate. See pages 348 and 349. *Photo: AP Photo/Leo Correa*



AAAS News & Notes 343
Science Careers 404

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. Copyright © 2023 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2411; Foreign postage extra: Air assist delivery: $107. 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.

CREDITS: (ILLUSTRATIONS, TOP TO BOTTOM) ROBERT NEUBECKER; PETER LODAHL

EDITORIAL

ChatGPT is fun, but not an author



H. Holden Thorp
Editor-in-Chief, *Science* journals.
hthorp@aaas.org;
@hholdenthorp

In less than 2 months, the artificial intelligence (AI) program ChatGPT has become a cultural sensation. It is freely accessible through a web portal created by the tool's developer, OpenAI. The program—which automatically creates text based on written prompts—is so popular that it's likely to be "at capacity right now" if you attempt to use it. When you do get through, ChatGPT provides endless entertainment. I asked it to rewrite the first scene of the classic American play *Death of a Salesman*, but to feature Princess Elsa from the animated movie *Frozen* as the main character instead of Willy Loman. The output was an amusing conversation in which Elsa—who has come home from a tough day of selling—is told by her son Happy, "Come on, Mom. You're Elsa from *Frozen*. You have ice powers and you're a queen. You're unstoppable." Mash-ups like this are certainly fun, but there are serious implications for generative AI programs like ChatGPT in science and academia.

ChatGPT (Generative Pretrained Transformer) was developed with a technique called Reinforcement Learning from Human Feedback to train the language model, enabling it to be very conversational. Nevertheless, as the website states, "ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers." Several examples show glaring mistakes that it can make, including referencing a scientific study that does not exist.

Many concerns relate to how ChatGPT will change education. It certainly can write essays about a range of topics. I gave it both an exam and a final project that I had assigned students in a class I taught on science denial at George Washington University. It did well finding factual answers, but the scholarly writing still has a long way to go. If anything, the implications for education may push academics to rethink their courses in innovative ways and give assignments that aren't easily solved by AI. That could be for the best.

More worrisome are the effects of ChatGPT on writing scientific papers. In a recent study, abstracts created by ChatGPT were submitted to academic reviewers, who only caught 63% of these fakes. That's a lot of AI-generated text that could find its way into the literature soon.

For years, authors at the *Science* family of journals have signed a license certifying that "the Work is an *original*" (italics added). For the *Science* journals, the word "original" is enough to signal that text written by ChatGPT is not acceptable: It is, after all, plagiarized from ChatGPT. Further, our authors certify that they themselves are accountable for the research in the paper. Still, to make matters explicit, we are now updating our license and Editorial Policies to specify that text generated by ChatGPT (or any other AI tools) cannot be used in the work, nor can figures, images, or graphics be the products of such tools. And an AI program cannot be an author. A violation of these policies will constitute scientific misconduct no different from altered images or plagiarism of existing works. Of course, there are many legitimate data sets (not the text of a paper) that are intentionally generated by AI in research papers, and these are not covered by this change.

Machines play an important role, but as tools for the people posing the hypotheses... and making sense of the results."

Most instances of scientific misconduct that the *Science* journals deal with occur because of an inadequate amount of human attention. Shortcuts are taken by using image manipulation programs such as Photoshop or by copying text from other sources. Altered images and copied text may go unnoticed because they receive too little scrutiny from each of the authors. On our end, errors happen when editors and reviewers don't listen to their inner skeptic or when we fail to focus sharply on the details. At a time when trust in science is eroding, it's important for scientists to recommit to careful and meticulous attention to details.

The scientific record is ultimately one of the human endeavor of struggling with important questions. Machines play an important role, but as tools for the people posing the hypotheses, designing the experiments, and making sense of the results. Ultimately the product must come from—and be expressed by—the wonderful computer in our heads.

–H. Holden Thorp

PHOTO: CAMERON DAVIDSON

10.1126/science.adg7879

NEWS



Protective equipment is required for researchers studying lethal viruses in a biosafety laboratory at Boston University.

PHOTO: CYDNEY SCOTT/BOSTON UNIVERSITY PHOTOGRAPHY

IN BRIEF

Edited by
Jeffrey Brainard

INFECTIOUS DISEASES

Biosafety group calls for clamping down on risky virus studies

The United States should expand oversight of "gain-of-function" research that tweaks risky viruses in ways that may make them more dangerous, concludes a draft report from two working groups of the National Science Advisory Board for Biosecurity (NSABB). A 2017 policy requires the National Institutes of Health (NIH) to seek reviews by its parent agency, the Department of Health and Human Services (HHS), for a subset of these studies: those "reasonably anticipated" to generate an "enhanced" version of a human pathogen that is likely to be both highly transmissible and highly virulent. The NSABB draft report calls for broadening this definition to studies that would create a pathogen likely to pose a severe public health threat and is at least moderately virulent, at least moderately transmissible, or both, such as SARS-CoV-2. Some experts say the proposed definition still gives NIH officials too much leeway to decide what should be reviewed; others worry it could sweep up routine studies important to public health. The full NSABB was to vote on 27 January on whether to forward the report to NIH and HHS for consideration.

Human geneticists apologize

RESEARCH ETHICS | The American Society of Human Genetics (ASHG) apologized this week for the participation of some of its early leaders in the eugenics movement and for other injustices committed by researchers in their field. Before the 1970s, many of its past presidents supported forced sterilization of people with "undesirable traits," and the society stayed silent about the use of genetics to justify discrimination against Black people, according to a report issued by ASHG. It wasn't until the early 1990s that the society formally opposed eugenics theories. Also during that decade, Arizona State University's Diabetes Project studied the genetic information of Havasupai tribe members without their consent, and ASHG remained uncritical. Historians had documented several of these cases, but their extent is "shocking, disturbing, and surprising," says Sarah Tishkoff, a University of Pennsylvania geneticist and member of an ASHG expert panel that guided the report, prepared by a contractor. To redress these harms, ASHG said it would strengthen diversity initiatives. It also will remove the names of all scientists from awards it bestows, pending a review of their actions.

Another HIV vaccine fails testing

IMMUNOLOGY | The beleaguered effort to develop an HIV vaccine has yielded yet another failure in the final stage of testing a candidate. An efficacy trial of a complicated vaccine regimen, developed by a Johnson & Johnson subsidiary in cooperation with the U.S. National Institutes of Health, stopped early because an interim analysis showed it offered no protection, sponsors announced

last week. The placebo-controlled study, dubbed Mosaico, assessed whether a series of six different shots could prevent HIV infection. Four shots delivered adenoviruses carrying a "mosaic" of genes from different HIV subtypes, and the final two contained two versions of HIV's surface protein. The trial involved 3900 men in Europe, South America, and the United States who have sex with men and with transgender people. A similar study in sub-Saharan African women that ended in 2021 also found the strategy offered no protection. No other experimental HIV vaccines are in or nearing efficacy trials.

African science leaders named

AGRICULTURE | Two institutes within CGIAR, an international organization that supports agricultural research, will soon be headed by African scientists for the first time. Cameroon-born Appolinaire Djikeng will direct the International Livestock Research Institute starting in April. It operates in Africa and Asia, with headquarters in Nairobi, Kenya, and Addis Ababa, Ethiopia, and an annual budget of about $80 million. Djikeng currently directs the Centre for Tropical Livestock Genetics and Health at the University of Edinburgh. Rwandan scientist Éliane Ubalijoro has been named CEO of the $100-million-a-year CIFOR-ICRAF, which funds research on deforestation, unsustainble food systems, and other topics. It includes two partners, the Center for International Forestry Research and World Agroforestry. Ubalijoro will also direct the latter, which is based in Nairobi. She comes from McGill University's Institute for the Study of International Development.

Convicted scientist avoids prison

RESEARCH SECURITY | An academic scientist in the United States convicted of failing to fully report his ties to a Chinese university will not be imprisoned or fined. In sentencing chemical engineer Franklin Tao last week, U.S. District Court Senior Judge Julie Robinson rejected the government's request that the former faculty member at the University of Kansas (KU), Lawrence, serve 30 months in prison and be fined $100,000. Robinson also chastised prosecutors for misunderstanding the culture of academic research, another setback to a Department of Justice initiative that has prosecuted some two dozen Chinese-born scientists. "There was no evidence of any conflict of interest … and this is not an espionage case," Robinson said in court. "What the government is claiming is a conflict of time [commitments]." Despite spending time in China, she said, Tao was also able to supervise research at KU "because apparently he is somebody who can work 70 or 80 hours a week consistently." KU put Tao, who was a tenured professor, on unpaid administrative leave after his August 2019 arrest and this month said he was no longer an employee.

War moves up Doomsday Clock

GLOBAL SECURITY | Russia's war against Ukraine helped prompt keepers of the iconic Doomsday Clock last week to advance it by 10 seconds, to 90 seconds before midnight. It's the clock's closest approach to "midnight"—which symbolizes humanity's self-annihilation—since the Bulletin of the Atomic Scientists began the annual exercise in 1947. Besides the Russia-Ukraine war, the group cited high risks from climate change, disinformation, and future pandemics.

Biden ups science board diversity

POLICY | President Joe Biden this month announced eight new members he plans to appoint to the body that oversees the National Science Foundation (NSF). The new class will give the National Science Board its most diverse makeup ever and reverse an emphasis on white, male appointees under former President Donald Trump. The board's new class includes seven women and one man, five of whom are scientists of color. In all, the 24-member board will feature 10 women, three Black scientists, and three Latino scientists. Board members serve 6-year terms, and one-third are replaced every 2 years.

Wolves swap sea otters into diet

ECOLOGY | Wolves that populated an island off Alaska's southeastern coast finished off nearly all the deer there, then turned to a surprising substitute: sea otters. The gray wolves on Pleasant Island attack the sea otters when they haul out onto the island's rocky shores to rest. It's a rare instance of a wolf population persisting without large terrestrial prey such as moose or deer, the research team reports this week in the *Proceedings of the National Academy of Sciences*. They said the finding highlights the unpredictable effects of predation among species that conservationists are trying to protect and reintroduce, including wolves and sea otters in Alaska.

GEOLOGY

AI helps find missed meteorites

Antarctica is famously good at preserving meteorites, burying the rocks in snow and ice until they resurface. They often become concentrated in regions of compacted "blue" ice that make up about 1% of the Antarctic surface. But finding the meteorites within those tracts has been an ad hoc affair. Now, researchers have a new tool for the hunt: an artificial intelligence algorithm that predicts their location. A team of Belgian researchers announced last week that it employed the machine-learning software to help discover five meteorites, including a massive, 7-kilogram specimen, in a blue ice region of East Antarctica not previously known to harbor them. The software, described in a study published last year in *Science Advances*, uses satellite data about known, meteorite-producing regions, such as slope, blue ice content, and temperature, and is trained to predict locations that have similar characteristics.

Scientists search an ice field in Antarctica for meteorites.



PHOTO: MARIA VALDES



IN DEPTH

INFLUENZA

Bird flu spread between mink is a 'warning bell'

Big outbreak at a Spanish farm reignites fears of an H5N1 influenza pandemic

By **Kai Kupferschmidt**

When mink at a big farm in Galicia, a region in northwestern Spain, started to die in October 2022, veterinarians initially thought the culprit might be SARS-CoV-2, which has struck mink farms in several other countries. But lab tests soon revealed something scarier: a deadly avian influenza virus named H5N1. Authorities immediately placed workers on the farm under quarantine restrictions. The more than 50,000 mink at the facility were killed and their carcasses destroyed.

None of the farm workers became infected. But the episode, described in a paper in *Eurosurveillance* last week, has reignited long-smoldering fears that H5N1 could trigger a human pandemic. The virus is not known to spread well between mammals; people almost always catch it from infected birds, not one another. But now, H5N1 appears to have spread through a densely packed mammalian population and gained at least one mutation that favors mammal-to-mammal spread. Virologists warn that H5N1, now rampaging through birds around the world, could invade other mink farms and become still more transmissible.

"This is incredibly concerning," says Tom Peacock, a virologist at Imperial College London. "This is a clear mechanism for an H5 pandemic to start." Isabella Monne, a veterinary researcher at the European Union's Reference Laboratory for Avian Influenza in Italy, where the samples from Spain were sequenced, calls the finding "a warning bell."

H5N1 was first detected at a goose farm in China in 1996. A big poultry outbreak in Hong Kong in 1997 led to the first documented human deaths and sparked the first pandemic fears. Around 2005, the virus spilled over into migratory birds, which have since spread it across the world in several big waves. A new variant named 2.3.4.4b that emerged in 2020 has spread faster and farther than any predecessor, dealing huge blows to the poultry industry in Europe and North America (*Science*, 26 August 2022, p. 912) before arriving in Central and South America in the fall of 2022. "It seems this virus is just more adapted to all birds than any others have been," says Richard Webby, an influenza researcher at St. Jude Children's Research Hospital.

Because the receptors the virus binds to in the upper airways of birds are less common in mammalian upper airways, H5N1 largely spares mammals. But this time around many mammalian species have become infected, including foxes, cats, ferrets, seals, and dolphins, presumably through contact with infected birds. On 17 January, Montana authorities said three juvenile grizzly bears euthanized in the fall after becoming very sick were infected with H5N1 as well. People have caught it, too. So far there have been six confirmed human infections in the current global wave, including one death.

There are some signs that 2.3.4.4b is less pathogenic in humans than earlier versions, which killed roughly half of those infected, says Thomas Mettenleiter, head of the Friedrich Loeffler Institute. "Of course that can be bad news, too, because it might make it easier for the virus to start spreading under the radar, giving it more opportunity to evolve," he says. The more often the virus infects mammals, the greater the risk, Webby says. "It's a numbers game."

There have been some past reports of avian influenza outbreaks on mink farms in China, but no clear evidence that the virus spread between the animals. In the Spanish outbreak, there seems to be little doubt it did. In theory, all of the sick animals could have picked up the virus from their feed, which included poultry byproducts, but H5N1 outbreaks have not been reported in the region where the poultry farms and slaughterhouses supplying the feed are located. And the virus spread from pen to pen as expected if it was transmitted between mink. The chain of infections might have begun after one animal caught a sick bird and pulled it into its cage, says Thijs Kuiken, a veterinary pathologist at Erasmus University Medical Center.

How readily the virus found in Spain

PHOTO: MADS CLAUS RASMUSSEN/RITZAU SCANPIX VIA AP

Outbreaks in farmed mink could give H5N1 an opportunity to further evolve and adapt to mammals.

might infect humans—or spread between them—is unknown. Sequenced virus samples from four mink show several changes compared with the bird virus, including T271A, a mutation in the gene for an enzyme, polymerase. The change, also seen in viral samples from other infected mammals, is known to help H5N1 better replicate in mammalian tissues. E627K, another worrisome mutation in the polymerase gene, has not emerged, however, and the gene for hemagglutinin—a protein on the viral surface that latches on to the host receptor—has not changed, Peacock says. "We may still have been lucky with this one."

Monne says her team and others are now studying the properties of the mink virus and the effects of the mutations it has accumulated. Among other things, they want to study how well the virus transmits through close contact between animals. "We are planning to also do aerosol transmission studies," she says.

The outbreak again puts the spotlight on the risks of mink farming. SARS-CoV-2, introduced at farms by humans, spread like wildfire among the animals but was also passed back to their caretakers, and researchers worried the mink industry might become a permanent source of infections and a breeding ground for genetic variants. The Netherlands, which had already decided to phase out mink farming by 2024 for ethical reasons, closed all remaining farms in 2021. Denmark culled all mink in the country in 2020, but a ban on mink farming expired at the beginning of this year.

The farms pose just as big of a threat when it comes to H5N1, Kuiken says. Most of the mammalian species infected with the virus so far are wild predators and scavengers feeding on infected birds—"solitary animals, or animals that live in small families," he says. They are unlikely to spread the virus far or infect humans. At mink farms, thousands of such solitary carnivores are forced to live together, creating ideal conditions for the avian virus to adapt to mammals. "It's a human construct," Kuiken says.

At the very least, biosafety measures on mink farms need to be tightened, Monne says. Farm workers should wear masks and take other measures to prevent infection, and the farms should reduce the risk of accidental H5N1 introductions. "They should be really carefully keeping the animals far away from wild birds." Peacock says perhaps it's time to just end mink farming. "That this is happening in Europe in this day and age, and after COVID-19, is doing my head in," he says. "It's a bit of an existential threat." ■

GRAPHIC: K. FRANKLIN/*SCIENCE*

WORKFORCE

Women at ocean science institute have half the lab space of men

Disparities exposed by "stunning" report from Scripps Institution of Oceanography are unlikely to be unique

By **Meredith Wadman**

Women constitute 26% of the scientists at the prestigious Scripps Institution of Oceanography (SIO) at the University of California, San Diego (UCSD), but only hold 17% of the space, according to an unprecedented report released last week.

Fifty-six women scientists at SIO have on average half as much research space and one-third the storage space of 157 of their male counterparts, according to the 95-page report by a task force of SIO faculty and staff and UCSD officials. (A few scientists who are nonbinary or did not confirm their gender were excluded from the analysis.) The 16 labs defined as "very large" all belong to men. Women also have less office space. And of 32 coveted storage containers in service yards on site—as opposed to at less convenient remote locations—31 are assigned to men.

The authors said the differences could not be "explained away" by funding, years at SIO, discipline, or research group size. "Our analysis points to the existence of widespread, institution-wide cultural barriers to gender equity within Scripps," they concluded.

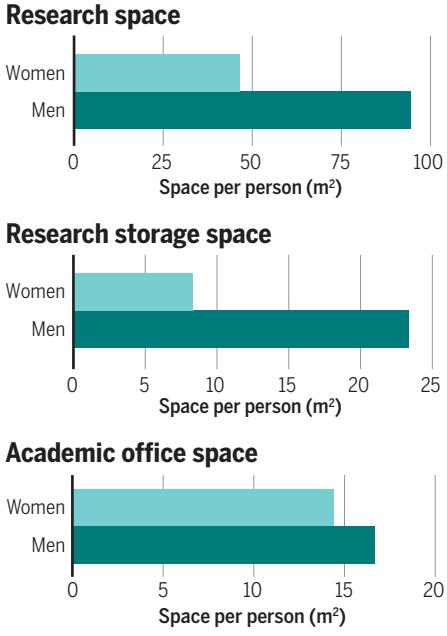
The report was commissioned in May 2022 by the university chancellor, executive vice chancellor, and SIO director after SIO faculty raised concerns. Its findings are likely to resonate in other institutions. American Geophysical Union President Lisa Graumlich, a paleoclimatologist at the University of Washington, Seattle, says that at major research universities she has visited nationwide, faculty from marginalized groups have told her they don't have enough space for their research and that space allocation policies lack accountability. She is "sadly not surprised" by the findings at SIO, she says.

The storied 120-year-old research center for ocean, earth, and atmospheric science, perched on a bluff above the Pacific Ocean, appears to be the first scientific institution to have conducted and released such an exhaustive statistical analysis of space allocation by gender. But its findings echo those of an investigation nearly 30 years ago led by Nancy Hopkins, now a biologist emerita at the Massachusetts Institute of Technology (MIT). In the early 1990s, under cover of dark, Hopkins measured every lab in the biology building there before spearheading a groundbreaking 1999 report on systematic discrimination against MIT faculty women. Hopkins calls the new results "stunning. … I looked at this thing and I thought, 'Oh my God, 30 years; I was doing this 30 years ago.' It has been written about and talked about and it's still happening."

The 1999 MIT report concluded that women there lacked space relative to men. But the data behind that finding were kept confidential. A 2000 gender equity review by the Woods Hole Oceanographic Institution found women scientists experienced a striking space deficit compared with their male peers as both advanced in their careers, but it did not examine possible confounders as the current study did.

Spaced out

Women at the Scripps Institution of Oceanography get short shrift compared with their male colleagues, according to a new report. The gaps can't be explained by differences in scientists' funding, years at the institution, or area of research.



When the authors of the new study corrected for variables such as funding, time at SIO, and discipline that might explain the stark differences in space assignments, they came up empty. As faculty gained more funding, space assignments for men grew at four times the rate that women's did. And as the size of their research groups grew, men's research space expanded at nearly double the rate of women's. The gender gaps persisted across research disciplines, meaning the clustering of men in a field that needs more space—say, oceangoing research versus computational studies—could not explain the discrepancies. Nor did research space track with the length of time a scientist had been at the institution, making it unlikely that a longer average tenure for men could explain some of the disparities.

The task force also illuminated dramatic differences in perceptions between men and women among 77 active faculty who responded to an anonymous survey. Asked whether they had sufficient space for their work, 42% of women said no, versus 6% of men. Only 10% of women found space assignments to be transparent versus 28% of men.



Former Scripps Institution of Oceanography geologist Jane Willenbring, in her lab there in 2019, praises Scripps for publicizing its new findings.

One contributor to the lopsided space allocations is a practice called "inheritance," the authors write. SIO policy requires that space be returned to the institution for reallocation when a faculty member moves, dies, or retires, but the policy has at times been ignored when a departing principal investigator simply assigns their space to an heir—a practice that has disproportionately benefited men, especially those with the largest labs.

Also contributing are emeritus faculty, 86% of them men, who hold nearly one-quarter of all space at SIO. Their capacious assignments are "difficult to comprehend," write Stefanie Lutz, an environmental hydrologist at Utrecht University, and Andrea Popp, a hydrologist at the Swedish Meteorological and Hydrological Institute, in an email to *Science*. The pair were lead authors on a 2019 global survey on the impacts of gender discrimination in earth and space sciences.

The new report, which UCSD posted on its website, "is exceptional in how thoroughly it was done—but also because [the UCSD administration] publicized it afterward. They could have just put it into a hole," says Jane Willenbring, a geologist at Stanford University who was an associate professor at SIO from 2016 to 2020.

UCSD Chancellor Pradeep Khosla wrote in a cover letter: "These findings do not reflect the values of our university." Khosla said he had directed SIO Director Margaret Leinen, who has been in the job 10 years, to chair a "Change Management" committee implementing the report's many corrective recommendations. Those include immediately identifying and reassigning available and underused space and "addressing the space assignments" of retired faculty to better serve those who aren't retired.

"[It's] gonna get fixed," says Victor Ferreira, a psychologist who is UCSD's associate vice chancellor for faculty diversity, equity, and inclusion and headed the task force that authored the report. "Everything I have seen including the fact that the public can download this report suggests that the university doesn't want to whitewash this problem."

It will take concerted corrective action to convince the skeptical. "Nancy Hopkins did all of this work and shone this light on how different it can be to be a woman in science than to be a man in science. And we have just learned nothing from that," Willenbring says.

Other research institutions may soon receive similar wake-up calls. One woman, a junior geoscientist at a major university who asked not to be identified for fear of career repercussions, says that in 2020, with COVID-19 protocols dictating the precise amount of space required per person in the lab, "suddenly there were spreadsheets flying around … and blueprints of the department." She soon generated a color-coded bar graph showing men at all career levels ahead of women in lab space per capita. "It just jumped out at me and I was like, 'Holy crap, this isn't good.'"

"This is still an ongoing problem for everyone at every level," adds a woman faculty member at SIO who asked not to be named because of the sensitivity of the issues. "This is not just geoscience or Scripps. This is all of STEM [science, technology, engineering, and math]." ■

PLANETARY SCIENCE

Meteorite results bode well for exo-Earths

Water and light elements are likely available right where rocky planets take shape

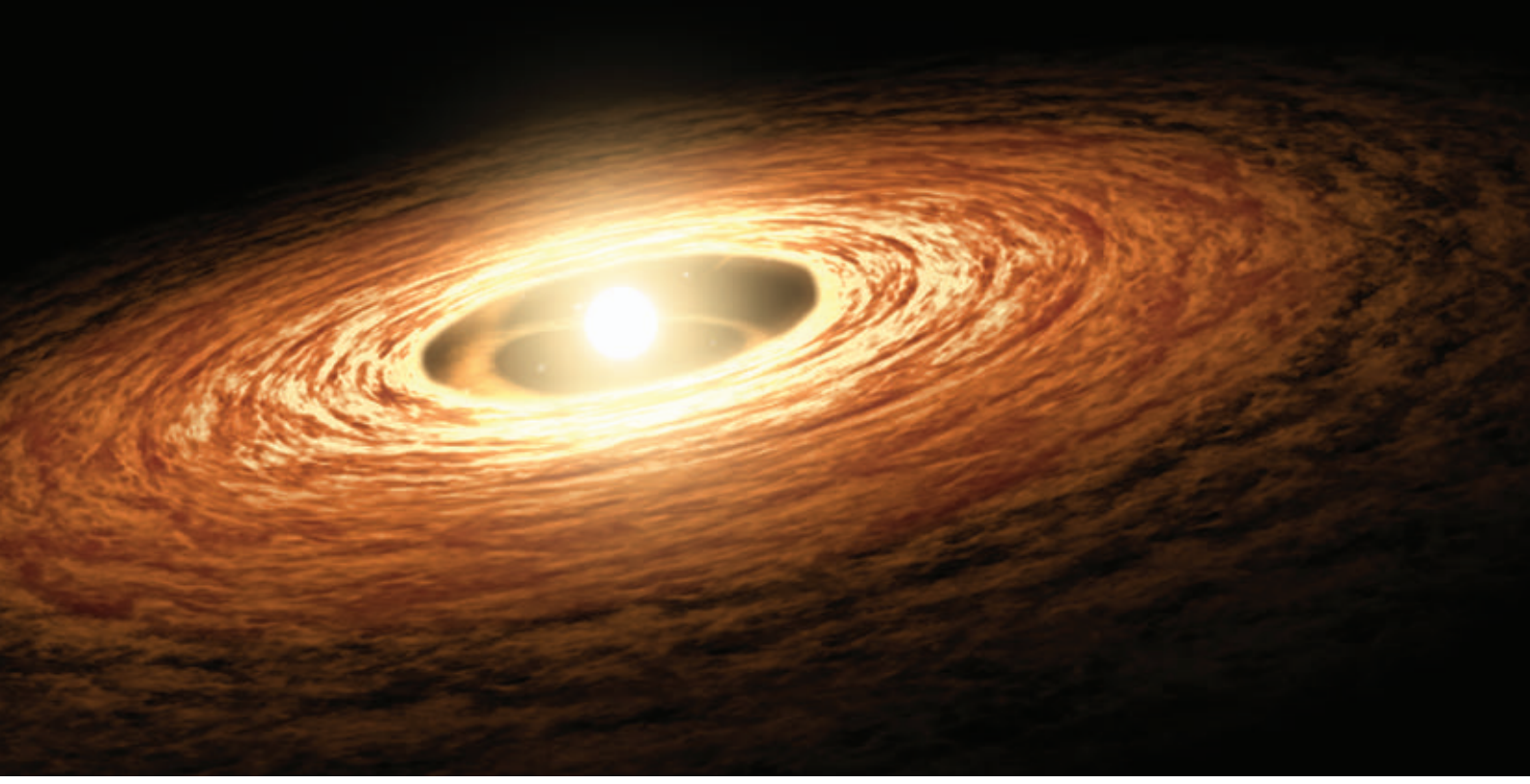
By **Paul Voosen**

How hard is it to give birth to an Earth? To assemble the right mix of rock, metal, and water, in a balmy spot not too far from a star? For a long time, planetary scientists have thought Earth was a lucky accident, enriched with water and lighter "volatile" elements—such as nitrogen and carbon—by asteroids that had strayed in from the outer edges of the early Solar System, where those materials were abundant. But a series of new studies, including two published today in *Science*, suggests all the ingredients were much closer at hand when Earth was born.

The findings, based on painstaking chemical analysis of meteorites, imply that planet-forming disks around other stars, too, should be well-stocked with the makings of wet, rocky planets that might be hospitable to life. "It makes the enrichment in volatile elements of a planet more generic," says Alessandro Morbidelli, a planetary scientist at the Côte d'Azur Observatory who wasn't part of the new work. Even if a young planet doesn't receive a delivery from the far reaches of the newborn planetary system, he says, "it doesn't change habitability."

Not that long ago, researchers thought the giant gas and dust disk that whirled around the early Sun more than 4 billion years ago had a fairly uniform composition. But that view was challenged by studies that tallied the ratios of certain isotopes found in the dozens of known species of meteorite. They indicated that the meteorites fell into two basic groups that likely originated in zones

IMAGES: (LEFT TO RIGHT) JANE WILLENBRING; NASA/JPL-CALTECH



A planet-forming disk around a young star (artist's concept) should readily give birth to habitable planets.

at different distances from the Sun. One group, known as carbonaceous chondrites, appears to have originated in the outer reaches of the early Solar System, beyond a proto-Jupiter and past the disk's "snowline," where temperatures cooled enough to allow water to freeze. In contrast, noncarbonaceous chondrites formed closer to the Sun. The isotopic signatures also suggested each zone was fed by material forged in different ancient and distant sources, such as supernovae and red giant stars (*Science*, 30 March 2018, p. 1451).

Until recently, scientists were only able to detect the early isotopic fingerprints in metals such as chromium, titanium, and molybdenum, which are durable enough to have resisted the heat of the newborn Sun. A close match between the isotope ratios found in noncarbonaceous chondrites and those found in the same metals on Earth suggested much of Earth's raw material came from the same nearby region as those meteorites.

But early searches for isotopic evidence that Earth's lighter volatile elements also originated nearby came up dry. "People just started thinking [the evidence] didn't exist," says Rayssa Martins, a doctoral student in geochemistry at Imperial College London (ICL). And so the traditional view of those elements' origin persisted: Many were assumed to come from a distant source, such as an outer region of the disk, where they might have condensed and then been pulled inward by the gravity of a forming planet, such as Jupiter.

Now, however, the meteorite divide has been detected in two moderately volatile elements, potassium and zinc. And the results suggest that much, but not all, of the planet's volatiles also came from the noncarbonaceous reservoir, says Nicole Nie, a planetary scientist at the California Institute of Technology and lead author of the potassium-focused paper. "This is a game changer for cosmochemistry."

The work was challenging, Nie says. Although a 2020 paper had identified what looked like ancient isotopic signatures in meteorite potassium, it only used two potassium isotopes, leaving out the much rarer potassium-40, whose signature in mass spectrometers is easy to confuse with those of calcium or argon. With only two isotopes, it was impossible to confirm that what the team saw reflected the chemical makeup of the primordial disk. And so Nie and her team measured all three potassium isotopes in 32 meteorites. They found that the potassium in the noncarbonaceous rocks showed isotopic patterns quite similar to those seen on Earth. "That was really surprising," she says. Together, the findings suggested some 80% of Earth's potassium came from nearby sources.

Three other teams have found a similar signal in the five stable isotopes of zinc. Two of the groups published their findings in *Icarus* last summer; work by the third, led by Martins, appears this week in *Science*. The findings complement each other, says Frédéric Moynier, a cosmochemist at the Paris Globe Institute of Physics and coauthor of one *Icarus* study. "I agree with everything in the [*Science*] paper, because it's very similar to our paper." Overall, it appears that half or more of Earth's zinc also came from the inner Solar System.

Other volatile elements probably had a similar origin, says Mark Rehkämper, a geochemist at ICL. "Zinc is not water. But where you have zinc, you will have more water." And although the newly formed inner Solar System was low in volatile elements overall, there was still enough to create a habitable world. "The water has been here almost from the beginning," Moynier says.

The hunt will now be on for additional volatile elements that exhibit the primordial fingerprints, says Thorsten Kleine, director of the Max Planck Institute for Solar System Research and co-author of one of the zinc papers. "We have just written the proposals to do that, to be honest," he says. Armed with enough data, especially for elements that are known to accumulate in a newborn planet at different stages of its growth, "you can do a detailed reconstruction of how the material that built the Earth changed over time."

That could help resolve another question that has nagged at planetary scientists for decades: how quickly the rocky planets were built. They may have formed slowly, over tens of millions of years, as smaller rocky bodies collided with each other, or much more quickly, as vast clumps of material collapsed. The isotopes could hold clues—and not just for our Solar System. As the new work makes clear, the recipe for Earth is unlikely to be a one-off. ■



In a 2016 experiment, researchers flooded an almond grove in a bid to restore California's groundwater.

PHOTO: UNIVERSITY OF CALIFORNIA, DAVIS

WATER RESEARCH

Can California's floods help recharge depleted aquifers?

Plans to drown orchards and farm fields to boost groundwater supplies get off to a slow start

By **Dan Charles**

The drenching storms that hit California in recent weeks represented a long-sought opportunity for Helen Dahlke, a groundwater hydrologist at the University of California, Davis. Dahlke has been studying ways to recharge the state's severely depleted groundwater by diverting swollen rivers into orchards and fields and letting the water seep deep into aquifers. But carrying out such plans requires heavy precipitation—which had been scarce in recent years.

This week, however, water managers began to turn theory into practice. In the Tulare Irrigation District, which supplies water to more than 200 farms south of Fresno, officials started diverting water from the San Joaquin River into 70 fields as well as specially constructed ponds. Each day, some 1.5 million cubic meters of water—roughly equivalent to 600 Olympic-size swimming pools—has been pouring onto the landscape. "We are in full [groundwater] recharge mode," Aaron Fukuda, the district's general manager, wrote in an email. Similar flooding is underway in the Madera Irrigation District north of Fresno.

Over the past decade, Dahlke's experiments with submerging small plots have suggested intentional flooding can replenish aquifers without damaging either groundwater quality or crops. But she says bureaucratic hurdles and organizational inertia have blocked widespread use of the practice—despite state laws and policies designed to encourage it.

"My frustration is growing!" Dahlke says. "This always looks so easy when you write these scientific papers, and give presentations, but to really implement [flooding] on a widespread scale is very hard." She and others hope this winter's intentional floods will encourage more of the state's water managers to embrace the practice.

California's farmers and others often extract far more water from aquifers than normally seeps in from the surface (*Science*, 17 April 2020, p. 230). The idea of using working farms to slow or reverse the trend was born in 2010, when independent hydrologist Philip Bachand and farmer Don Cameron flooded some of Cameron's vineyards. The vines thrived, and the water replenished the aquifers beneath Cameron's land.

Four years later, California adopted a landmark law, the Sustainable Groundwater Management Act (SGMA), that promotes the practice. It requires farmers to treat aquifers like bank accounts, clamping down on overdrafts but also allowing those who deposit water into them to make bigger withdrawals later.

The most catastrophically depleted aquifers lie in the San Joaquin Valley, the nation's largest single source of tree nuts, fruit, and vegetables. In places, groundwater extraction has caused the land to sink by several meters, and declining runoff from the Sierra Nevada means growers can no longer depend on a steady supply of river water. In this region, Dahlke says, capturing water during wet years and storing it underground for later use will be a matter of survival. The looming shortage "is just getting really scary," she says.

But several obstacles have stood in the way of recharge projects, experts say. Some districts need state permits and getting them is time-consuming. The SGMA's limits on extraction are only kicking in now, so farmers haven't had much incentive to spend the money required to flood their fields. "If you've chosen to somewhat ignore this law, you've been able to," says Sarah Woolf, a water consultant and farmer.

Still, the recent recharge efforts are prompting new interest. In the Madera Irrigation District, General Manager Thomas Greci says farmers seem increasingly open to drenching their fields. "I have been shocked to see the number of growers coming in to sign up to take this water," he says. And other irrigation districts have been calling, asking how it's done, says Dina Nolan, the district's assistant general manager. "It was, frankly, quite shocking to me," she says. "I was like, 'You've never promoted this?'"

California's high waters are now receding, but the opportunity to capture runoff will likely continue through the spring as a hefty mountain snowpack melts. Many farmers, however, won't be inclined to drown their fields when it's time to plant or pollinate their crops. "Only certain crops are compatible [with flooding] at that time of year," says Daniel Mountjoy, director of resource stewardship with the nonprofit Sustainable Conservation.

When all is said and done, Dahlke estimates this year's intentional floods will counterbalance less than 10% of the San Joaquin Valley's typical annual groundwater deficit. But she hopes the experience will prepare the state to do better when the next deluge arrives. With that in mind, she's hoping to soon launch a study aimed at identifying easier ways of using California's extensive irrigation infrastructure to steer 1 trillion liters of floodwater into the state's aquifers. The goal, she says, is "to go more large-scale." ■

Dan Charles is a journalist in Washington, D.C.

ENVIRONMENT

Fukushima wastewater release set to start soon

Japan, facing broad opposition, says dump will pose no risk to Pacific Ocean life or humans

By **Dennis Normile**, *in Tokyo*

The Japanese government is pushing ahead with its plan to release 1.3 million tons of radioactive water from the defunct Fukushima Daiichi Nuclear Power Plant into the Pacific Ocean. The release could begin as early as this spring or summer, according to materials distributed at a 13 January ministerial meeting. But it has stirred broad opposition—from Japan's fishing industry and consumers, countries throughout the Asia-Pacific region, and some marine scientists.

The Tokyo Electric Power Company (TEPCO), which owns the power station, says it is running out of space to store the water on land. Radioactivity levels in the discharged water will be too low to pose a risk to marine life or humans, TEPCO says, and its plan has the blessing of the International Atomic Energy Agency (IAEA). At a 20 January briefing here, IAEA nuclear safety official Gustavo Caruso, who heads a special agency task force on Fukushima, said Japan's Nuclear Regulation Authority (NRA) has procedures in place to ensure the discharge meets international safety standards.

But critics say the risks haven't been studied in enough detail. TEPCO's assurances are "not supported by the quantity and quality of the data," says oceanographer Ken Buesseler of the Woods Hole Oceanographic Institution. "We need more information." The release would set an awful precedent, says Robert Richmond, a marine biologist at the University of Hawaii, Manoa: "There is a strong consensus internationally that continued use of the ocean for dumping waste is simply not sustainable."

Ever since the 2011 Tohoku earthquake and tsunami devastated the Fukushima power plant, crews have continuously pumped water through the wrecked reactors to cool the nuclear fuel, much of which melted. The cooling water picks up radionuclides, many of which are then captured by a specially developed filtering process. But tritium, which is chemically identical to the ordinary hydrogen in water, slips through the system. For now the water is stored in more than 1000 tanks on the plant grounds.

Tritium gives off only low-energy beta particles believed to pose minimal risks for marine life and humans. And TEPCO plans to dilute the waste with massive amounts of seawater to reduce tritium levels below regulatory standards for drinking water before dumping it into the ocean through a pipe extending 1 kilometer offshore.

A level safe for drinking sounds reassuring, but is still thousands of times higher than the natural level in seawater, Richmond says. And the water will be discharged at a single point for decades, so tritium, which can be bound into animal and plant tissue, and other radioactive isotopes may still accumulate in marine organisms and work their way up the food chain to fish and humans. "This has real implications for ocean life and the human lives that are tied to our oceans now and into the future," Richmond says.

Other radionuclides may also be slipping through, says Ferenc Dalnoki-Veress, an expert in low concentration radioactivity measurements at the Middlebury Institute of International Studies at Monterey. "What's in the water? We don't really know," he says. TEPCO has sampled small amounts of water from just one-quarter of the tanks, he says, and measured concentrations of tritium and only a limited number of other radionuclides. Strontium-9 and cesium-137, radioactive products of nuclear fission, have turned up in wildly varying concentrations, raising questions about how well the filtration system works. (TEPCO maintains that further filtering will capture more of the radionuclides.)

Japan's neighbors share the concerns. "There should be no discharge until all parties verify through scientific means that [it] is safe," Henry Puna, secretary general of the Pacific Islands Forum, a group of 16 countries, said at a recent seminar. The U.S. National Association of Marine Laboratories also opposes the plans, citing "a lack of adequate and accurate scientific data supporting Japan's assertion of safety."

Scientists have put forward alternatives. Richmond sees an opportunity for bioremediation, by pumping the wastewater through tanks full of oyster species that consume plankton and incorporate radionuclides into their shells. If other radionuclides are removed, the contaminated water could simply be stored for 40 to 60 years, since tritium has a half-life of only about 12 years. Or the water could be used to make concrete, from which tritium's beta particles are unable to escape.

But TEPCO is pressing ahead. NRA will monitor the entire operation, as will IAEA and at least one third party commissioned by TEPCO. Independent groups will also be watching. A group led by Núria Casacuberta Arola, an oceanographer at ETH Zürich, collected samples off the coast of Fukushima in November 2022 that will provide baseline concentrations of tritium and other radionuclides, with which any changes can be compared. But more studies should be done before the releases begin, Richmond says: "Monitoring does not prevent problems from occurring, it identifies when problems occur." ■



Radioactive wastewater is now stored in more than 1000 tanks near Fukushima's destroyed reactors.

PHOTO: *THE ASAHI SHIMBUN*/GETTY IMAGES



READY, SET, SHARE!

As funders roll out new requirements for making data freely available, researchers weigh costs and benefits

By **Jocelyn Kaiser** *and* **Jeffrey Brainard**

Physiologist Alejandro Caicedo of the University of Miami Miller School of Medicine is preparing a grant proposal to the U.S. National Institutes of Health (NIH). He is feeling unusually stressed because of a new requirement that takes effect this week. Along with his research idea, to study why islet cells in the pancreas stop making insulin in people with diabetes, he will be required to submit a plan for managing the data the project produces and sharing them in public repositories.

For his lab, that's a daunting task. Unlike neuroscience or genomics, Caicedo's field has no common platforms or standards for storing and sharing the kinds of data his lab generates, such as videos of pancreatic islet cells responding to a glucose stimulus. The "humongous" raw imaging files are currently stored in an on-campus database, notes Julia Panzer, a postdoctoral researcher in the lab. To protect patient privacy, the database is secured and not designed to provide access to outsiders. Sharing the data will mean uploading them somewhere else.

Caicedo supports the new NIH policy, acknowledging that "science will be so much more powerful" if data are freely shared. But he says his field isn't ready. And he's worried about the burden the new mandate will impose on his postdocs and graduate students. He can't afford to hire a data manager for his eight-person lab with his $600,000 in NIH grants, he says. "It's a very limited budget for a lot of people."

In the years ahead, many researchers will be struggling with similar issues. By 2025, new U.S. requirements for data sharing will extend beyond biomedical research to encompass researchers across all scientific disciplines who receive federal research funding. Some funders in the European Union and China have also enacted data-sharing requirements. The new U.S. moves are feeding hopes that a worldwide movement toward increased sharing is in the offing. Supporters think it could speed the pace and reliability of science.

Some scientists may only need to make a few adjustments to comply with the policies. That's because data sharing is already common in fields such as protein crystallography and astronomy. But in other fields the task could be weighty, because sharing is often an afterthought. For example, a study involving 7750 medical research papers found that just 9% of those published from 2015 to 2020 promised to make their data publicly available, and authors of just 3% actually shared, says lead author Daniel Hamilton of the University of Melbourne, who described the finding at the International Congress on Peer Review and Scientific Publication in September 2022. Even when authors promise to share their data, they often fail to follow through. Out of 21,000 journal articles that included data-sharing plans, a study published in *PLOS ONE* in 2020 found, fewer than 21% provided links to the repository storing the data.

Journals and funders, too, have a mixed record when it comes to supporting data sharing. Research presented at the September 2022 peer-review congress found only about half of the 110 largest public, corporate, and philanthropic funders of health research around the world recommend or require grantees to share data.

"Health research is the field where the ethical obligation to share data is the highest," says Aidan Tan, a clinician-researcher at the University of Sydney who led the study. "People volunteer in clinical trials and put themselves at risk to advance medical research and ultimately improve human health."

Across many fields of science, researchers' support for sharing data has increased during the past decade, surveys show. But given the potential cost and complexity, many are apprehensive about the NIH policy, and other requirements to follow. "How we get there is pretty messy right now," says Parker Antin, a developmental biologist and associate vice president for research at the University of Arizona. "I'm really not sure whether the total return will justify the cost. But I don't know of any other way to find out than trying to do it."

Science offers this guide as researchers prepare to plunge in.

ILLUSTRATION: DAVIDE BONAZZI/SALZMANART

Why share scientific data?

Scientists who share data can speed up and improve science and advance their own careers, advocates say.

Growing efforts in psychology, cancer research, and other fields to reproduce published studies, for example, depend on access to the underlying data. Such access can help resolve whether an apparent error in a paper resulted from an honest mistake—or faked data. In 2020, for example, a high-profile study claiming that COVID-19 patients given the antimalarial drug hydroxychloroquine were at increased risk of death fell apart after Surgisphere, the company that claimed to have provided the underlying data set, couldn't produce it. (Studies purporting to show benefits from the drug, touted by then-President Donald Trump, proved equally problematic.)

Sharing data could also help curtail duplicative efforts to collect them. That could save time and money for smaller labs in particular, says Crystal Rogers, a cell and developmental biologist at the University of California (UC), Davis. "Maybe this policy will even the playing field," she says. "It will democratize opportunities."



Existing data can help researchers generate hypotheses, design clinical trials, and teach. And by pooling smaller data sets, scientists can conduct meta-analyses that can produce robust or intriguing findings, says Maryann Martone, a neuroscientist at UC San Francisco. She points to a study that gathered raw data from an array of animal studies conducted in the 1990s on treatments for spinal cord injury. The results from the individual studies were inconsistent and never published. But a 2021 analysis of pooled data from 1125 animals produced a significant correlation: Animals with blood pressure levels within a certain window during spine surgery fared better, a finding that held up in a clinical study. "There's real gold in these small data sets, if you can put them together," Martone says.

For the researchers who share their data, one proven reward is increased citations to papers for which data are provided. Papers that provided a link to data gained 25% more citations on average than those that did not, according to a 2020 study of more than 50,000 articles in the PLOS and BMC journals.

Even as more funders expect grantees to provide data, a lack of professional rewards may be responsible for widespread noncompliance. Sharing typically doesn't count for much in tenure and promotion reviews, for example. Academic institutions should encourage departments to develop policies for providing such rewards, according to a 2021 joint report from the Association of American Universities and the Association of Public and Land-grant Universities.

It may be hard to overcome fears that researchers who share data won't get proper credit from others—or may even get scooped. "How do you make sure that somebody doesn't grab that data and publish it as their own in some minor journal?" worries cancer physician-scientist Jan Grimm of Memorial Sloan Kettering Cancer Center. Advocates for data sharing have called for publishers to discourage such behavior by requiring authors who use data generated by other scientists to name them as "data authors."

Scientists may come to see data sharing as a useful burden, like peer review, says Tim Vines, founder of a data search tool called DataSeer. "Peer review is very annoying, but many people say: 'It improves my manuscripts.' Researchers accept that. We need to bring [data] sharing to that level."

How are data-sharing policies changing?

Many U.S. funders already have sharing policies. NIH has been a leader of such efforts, rolling out a 1996 policy for its grantees in human genome sequencing and expanding it in 2003 to cover all large projects. Now, the agency is extending its rule to cover all of its research grants.



NIH's new policy "strongly encourages" researchers to deposit project data in repositories where other researchers have free access. The data should be "of sufficient quality to validate and replicate research findings," the policy says. Data should be deposited when a journal article about them is published or the grant ends, whichever comes first. And the policy extends to unpublished findings, including negative results.

"We really wanted to catalyze the research community through a more ubiquitous data-sharing policy," says Lyric Jorgenson, acting director of NIH's Office of Science Policy, who oversaw development of the policy.

NIH's push could test the kinds of changes that all federally funded researchers will need to make by December 2025, when a revised data-sharing policy announced in August 2022 by the White House Office of Science and Technology Policy (OSTP) takes effect. (The new policy also drew attention for requiring that journal articles be free to access when published.)

Like NIH's policy, OSTP's requires all data "underlying" peer-reviewed scholarly papers be made publicly available for free when papers are published (although it allows exceptions, including if taking that step is too costly). The policy is significantly more stringent than the current requirement at a key agency it covers, the National Science Foundation (NSF), which only requires data sharing within "a reasonable time period." NSF and other U.S. research-funding agencies are expected to propose details this year and next about how they will implement OSTP's policy.

How do I comply?

NIH's policy requires investigators submitting a grant proposal to include a two-page data-management plan listing the types of data they will produce, the software or tools needed to use the data, and the publicly accessible repositories where they will be stored. When submitting the data, researchers will need to include "metadata," or details of how the data were collected. And they may need to reformat data to fit a repository's standards. These steps are meant to make the data comply with international guidelines called the FAIR principles, which stands for "findable, accessible, interoperable, and reusable."

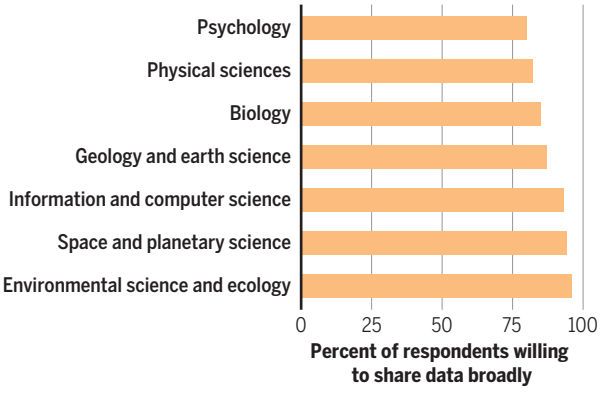


Supporters of sharing also call for something encouraged, but not required by NIH: choosing repositories that attach digital object identifiers (DOIs) to data sets that are different from those used to identify the associated papers. The DOIs—unique, permanent serial numbers—will make it easier for other researchers to find relevant data. (Authors and journals must also properly format a manuscript's references to associated data for them to be discoverable by search tools.) DOIs will also identify each data set as an independent scholarly contribution, enabling researchers to claim credit for generating and sharing the data.

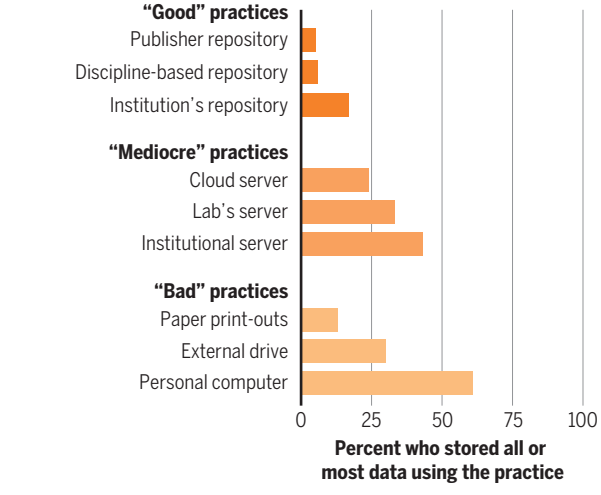
ILLUSTRATIONS: A. MASTIN/*SCIENCE*

Data sharing, by the numbers

Scientists express interest in sharing their data ...
Across fields, most scientists like the idea of sharing data, according to a 2017–18 survey of more than 2000 respondents from multiple countries. (Selected categories are shown.)



... but few do it right, if at all.
When researchers do share data, many don't use practices that promote long-term curation and sharing with other researchers. Analysts say big changes in research culture are needed, and more training and guidance could help. (Survey respondents could report more than one type of storage method.)



What would motivate researchers to share data?
Professional rewards rank highest, but researchers also see public benefits, according to a survey of more than 6000 respondents in 2022. (Selected motivations are shown, and respondents could choose more than one reason.)



Some researchers who recently began to share more data under existing NIH policies say the process can be very time-consuming. Relabeling, reformatting, and otherwise preparing all the underlying data collected by co-authors on a paper can take half a day, says Florian Krammer, a virologist at the Icahn School of Medicine at Mount Sinai—work he typically does on a weekend. His data manager needs another full day to upload the data to databases. "I think a lot of people don't realize how much work it is," he says.

Others point out that if researchers develop plans for data sharing at the start of a project, the costs may go down. "The time decreases the better managed your lab is, because things are documented from the get-go instead of at the very end," UC San Diego's Martone says. Stacey Schultz-Cherry, a virologist at St. Jude Children's Research Hospital, puts it this way: "We're all going to grumble, but in the long run it's really going to benefit science."

What about sensitive data that are difficult to share?

Biomedical research often involves human subjects who may not have agreed to having their data shared, even if they have been stripped of identifying information. NIH's policy allows exceptions for such data. But the agency expects that, when possible, consent forms for new studies will ask participants to agree to share their deidentified data.



Although some institutional ethics boards have opposed such broad consent, "I think there's an understanding that this is what the community is moving towards," says physician Ida Sim of UC San Francisco. She is a co-founder of Vivli, a repository that some institutions plan to use to share participant-level data from clinical trials.

Clinical trial researchers are known for keeping their data under wraps because they're concerned they won't be analyzed properly, or they're still writing papers. Many already ignore a 2016 NIH rule requiring that summary data be posted in the federal ClinicalTrials.gov database no later than 1 year after the trial's primary completion date. But the NIH data-sharing policy is already "being taken as a serious mandate" by those scientists, Sim says. "I am pleased with how much of a cultural change this has catalyzed."

How will policies be enforced?

Because biomedical disciplines create and use data differently, NIH says it chose to provide flexibility by not packing its policy with detailed requirements.



In particular, it does not specify how much data researchers must share from a given data set. Do they need to deposit an entire video of dividing cells or a molecular marker infiltrating a tumor, which could be gigabytes of data, or just the still images presented in papers? "Many of us do not fully understand at what level, from raw to fully processed and grouped, NIH expects data to be shared," says cardiovascular disease researcher Curt Sigmund of the

CREDITS: (ILLUSTRATIONS) A. MASTIN/*SCIENCE*; (GRAPHICS) K. FRANKLIN/*SCIENCE*; (DATA, TOP TO BOTTOM) C. TENOPIR *ET AL. PLOS ONE*, 15(3), E0229003(2020) (2); G. GOODEY *ET AL., THE STATE OF OPEN DATA 2022*; DIGITAL SCIENCE REPORT. HTTPS://DOI.ORG/10.6084/M9.FIGSHARE.21276984.V5

Medical College of Wisconsin. The answer, NIH's Jorgenson says, is that each discipline will need to work out the "granularity" required to reproduce a paper's findings.

In practice, NIH program managers will review an investigator's sharing plan when a grant proposal is submitted and check progress reports to be sure the plan is being followed. The agency could terminate a grant for noncompliance, although that rarely happens for violations of other NIH policies. But those who don't share data could be barred from receiving a new grant, Jorgenson says. Achieving the data-sharing policy's goals will likely be achieved "in stages and steps," she adds. "We did not want to set the bar so high that we create disincentives for anyone to participate."

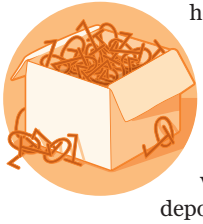
Many funders and journals have struggled to enforce their own sharing requirements. Confirming whether authors shared all data supporting an article can require a close, time-consuming examination, Vines says. Publishers receive no extra revenue for the added effort.

To avoid data sharing that is incomplete or poorly done, funders and institutions may need to not only threaten researchers with sticks, but also offer them carrots, in the form of technical support and training, says Dylan Ruediger, a project manager at Ithaka S+R, a higher education research and consulting organization. He managed an NSF-funded project that brought together interdisciplinary teams of researchers in fields as disparate as agronomy, nuclear imaging, and polar science to examine barriers to data sharing.

"Complying with mandates to deposit data is not the same thing as creating an ecosystem that's really well adapted to help researchers reuse data," Ruediger says. "That's a very different kind of challenge."

How and where do I store data?

Currently, many researchers store data primarily on their personal computers (see graphic, p. 324). Sharing data will mean shifting them to one of several possible homes: a repository at the researcher's institution; a discipline-based one, such as OpenNeuro, which holds brain-imaging data, or NIH's ImmPort, which stores immunology data; or a general repository, such as figshare or Zenodo. Many repositories will need improvements to make it easier to deposit, find, and retrieve data, experts say.



To help navigate this new terrain, some universities are beefing up staff who can help, such as IT specialists and librarians who specialize in data. "We're reprioritizing some of the things that we're doing in the library to accommodate these requests," says Vicki Coleman, dean of library services at North Carolina A&T State University, a historically Black research institution. She says the library will shift staffing away from its traditional reference desk—a trend underway at other universities as well.

These data experts often have clever ways to adapt commonly used information-management tools to the needs of specific research fields. Many universities, for example, now offer faculty members training in using Jupyter Notebooks, an open-source web application designed to make it easier to share data. The extra staffing and training should address a concern Ruediger found among participants in his project to encourage data sharing: "a sense that the challenges they were facing were unique and idiosyncratic to them."

What are the costs, and who will pay?

Scientists say it can be difficult to estimate the cost of cleaning and preparing data for use outside their team. For example, Krammer of Mount Sinai estimates data sharing will eat up at least 10% of his funding. Hiring a data manager might cost $100,000 per year, although not all labs will need one.



NIH says researchers applying for a grant can add costs for data managers, staff time to prepare data, and repository fees. But because NIH has a strict dollar limit for many grants, data-sharing costs may cut into the funds available for research. "If you're loading up your grant budget with data-sharing and management costs, is that going to take away from the funds for doing the science?" says David Kennedy, vice president of the Council on Governmental Relations, which represents major research universities.

Universities, for their part, will have to pay for campus-wide services supporting data sharing, such as librarians and subscriptions to repositories. Institutions can bill these "indirect" costs to the overhead that funders provide with grants. But those reimbursements are capped. Although universities have long tapped their own revenues to help cover the indirect costs of research, some worry data sharing will become another "unfunded mandate" from the federal government.

The costs per institution will exceed $1 million a year, split between overhead and investigators' budgets, according to an initial analysis based on a survey of 34 Council on Governmental Relations members. That could be a special burden for smaller institutions, Kennedy says. That is "a huge concern," Jorgenson acknowledges. "We do not want to exacerbate inequities in the funding structure."

Another challenge to be solved: Even the largest repositories are still looking for sustainable business models. Discipline-specific ones are typically supported by grants for individual projects that don't assure funding after the grant ends. NIH's and OSTP's policies don't spell out for how long data must be stored and shared; Jorgenson says the agency "will be collecting lots of information" to inform a more specific policy on this.

Will broad data sharing be worth the effort?

Skeptics say the benefits are yet to be demonstrated. Krammer says funders should collect and analyze data about whether the new push is producing the intended effects. "There needs to be an evaluation after 2 years, 5 years, to look at what type of data is [re]used, and for what type of data it doesn't seem to make sense," he says.



Supporters of data sharing agree—and think the results will bear them out. "We need some real demonstrations of how this level of data sharing can drive the discovery engine," UC San Francisco's Sim says. "I don't think we're there yet. But it's kind of like everyone's hopped in the car, and we're starting the engine." ■

ILLUSTRATIONS: A. MASTIN/*SCIENCE*



POLICY FORUM

ANIMAL BEHAVIOR

Farm animal welfare: Beyond "natural" behavior

An animal-centered view guided by what animals value could improve welfare on farms

By **Marian Stamp Dawkins**

Globally, more than 78 billion land animals are reared for human consumption each year (more than 70 billion chickens, with pigs and cattle making up much of the rest) (*1*). The trend is growing, but so is public concern about the welfare of these animals, particularly in intensive farming systems. In response to public concerns about the lives of farm animals, requirements for improved animal welfare now widely appear in legislation, in farm assurance schemes, and as an important ESG (environment, social, and governance) goal for food-producing companies, giving animal welfare a higher priority than it has ever had before. However, requirements that prioritize animals' ability to exercise "natural behavior" are often imposed without showing that they actually improve welfare from the animals' point of view. Without evidence to inform policies and practices, industry may be wasting money and misleading the public—all without any genuine improvement in animal welfare.

Approaches to animal welfare have included the Five Freedoms (*2*), the Five Domains (*3*), and the Four Principles of the European Union's Welfare Quality Assessment. All put good physical health (particularly absence of disease, injury, hunger, and thirst) as necessary components of welfare. But they all also agree that there is more to good welfare than physical health, and it is in trying to capture that extra component of welfare—what might be called the mental as well as the physical health of an animal—that the real problem lies.

The food industry has addressed this problem by placing increasing emphasis on one particular welfare metric: natural behavior. Natural behavior is behavior shown by animals living where their ancestors evolved or at least in man-made environments that allow them similar freedom of movement. The welfare of more confined members of the same species, such as those living in intensive farming systems, is then judged by the extent to which they too are able to show this natural behavior [for example, (*4*)].

The critical assumption here is that natural behavior is necessary for good welfare and therefore that welfare standards on a

Department of Biology, University of Oxford, Oxford, UK.
Email: marian.dawkins@biology.ox.ac.uk

PHOTO: ANGEL GARCIA/BLOOMBERG/GETTY IMAGES



A cow uses a scratching brush at La Fageda farming cooperative in Santa Pau, Spain, in August 2022.

farm can be judged by the extent to which animals living there show this natural behavior. But this assumption has been repeatedly challenged—for example, on the grounds that some natural behavior, such as predation, actually reduces welfare (*5*, *6*). Given the priority now being given by the food industry and the public to natural behavior [for example, (*4*)], the objections that have been raised need to be taken more seriously by those responsible for actually introducing changes to farm animal management. Though useful, natural behavior is only a starting point for assessing welfare; overemphasizing it may mean that animal welfare is not being improved as much as it could be.

THE RISE OF NATURAL BEHAVIOR AS A WELFARE METRIC

"Natural" is an emotionally loaded word, exploited by countless advertisers wanting to present their products as desirable because they are natural as opposed to artificial. So "natural" and "behavior" form a combination that finds approval almost without question. It is therefore not surprising to find that public opinion supports the idea that animals should be free from the restriction of cages and able to behave naturally (*7*).

The idea of natural behavior as a necessary part of good welfare really took hold with the publication in 1979 of the Five Freedoms (*2*), five simple statements about what was necessary for good animal welfare. The first three of the Five Freedoms address the conditions necessary for physical health—freedom from hunger, physical discomfort, and injury. The last two refer to mental health—freedom from stress and, importantly, freedom "to perform most normal patterns of behavior" ("normal" rather than "natural" being used to allow that domesticated farm animals may behave differently from their wild ancestors but still have the same need for freedom of movement).

Natural behavior was given this prominent place in welfare assessment because it was believed by many at the time that animals had natural instinctive urges that had to be satisfied if animals were not to be frustrated. "A very large part of animal behaviour is basically determined by instinctive or innate abilities, proclivities, and dispositions" wrote W. H. Thorpe in one of the key papers used in drawing up the Five Freedoms [(*8*), p. 73]. "Suppression of these instinctive appetites can give rise to evidence of prolonged and intense emotional disturbances" [(*8*), p. 79]. Natural behavior and welfare thus seemed inextricably linked.

The Five Freedoms were subsequently taken up and used around the world as the basis for animal welfare legislation and as part of industry practice. With the current interest in animal welfare, their influence today is perhaps greater than ever, and they are increasingly found, often in their original wording, in the sustainability and ethical goals of food producers, food retailers, and outlets across the world. For example, an international food outlet gives as one of their animal welfare priorities "Providing enrichments that support natural behavior" and states, in applying this to chickens, "We are working with supply chain partners to ensure housing environments that promote natural behaviors such as pecking, perching, and dust-bathing" (*9*).

Natural behavior thus seems to be the ideal welfare metric. It is based on the animals' own behavior, can be measured and even quantified, has political clout around the world, and above all has public opinion on its side. All of this makes it all too easy to overlook the very real problems there are with uncritically linking natural behavior to welfare. Natural behavior, if used on its own, may not improve animal welfare at all.

THE FLAWS WITH NATURAL BEHAVIOR AS A WELFARE METRIC

One obvious problem with natural behavior is that a life in the wild can be brutal. Wild animals in their natural environments are subject to constant threats to their survival. Even in populations completely unaffected by humans, death by predation, disease, malnutrition, attacks by other animals, parasites, and adverse weather happen all the time and results in many wild animals having very short lives. Captive animals are often healthier and live longer than their wild counterparts (*10*). At least in captivity, an injured animal can be recognized and treated quickly, whereas a wild one can be left to a slow and lingering death or eaten while still alive. Many a natural environment would not pass muster if judged by the standards of the Five Freedoms.

Just as living in a natural environment does not guarantee good welfare, neither does performing natural or species-typical behavior (*5*, *6*). Being chased by a predator, for example, is entirely natural for many wild animals but is also something they will choose to avoid if given the opportunity. Here is the real weakness of natural behavior as a welfare criterion. It does not take into account what the animals themselves want (*5*, *6*). What an animal does—the behavior that it shows—simply reflects the state it is in at any moment, and that could be a state of poor welfare that it was attempting to change. For example, a wild animal could be hiding from danger because it was fearful or foraging for food because it was desperately hungry. In other words, the natural behavior an animal is showing can be a sign that its current state of welfare is poor, and the animal is attempting to rectify that.

WHAT ANIMALS VALUE AS A WELFARE METRIC

The past 40 years have seen major advances in our understanding of the behavior of animals: what stimuli are important to them, what motivates them, and crucially, what are the effects of not being able to perform certain behaviors. The current view is that most animals, including the birds and mammals that form part of modern farming, have some behavior that can be described as innate or unlearnt, but they are not automata rigidly programmed to carry out a fixed repertoire of natural behaviors and doomed to be frustrated if they cannot perform every single one exactly the way and in exactly the proportions shown by their wild ancestors. On the contrary, natural selection has made animals much more flexible than this by acting

more on what animals value (what they find rewarding or aversive) than on the specific behaviors they might perform (*11*).

What animals have evolved to do is not to "perform a behavior" but to be rewarded or punished by circumstances that are the usual result of such behavior, which gives them much greater range of behavioral options (*11*). For example, an animal that has been selected to find a sweet taste pleasant can learn to find food in environments that may be completely different from anything its ancestors ever experienced, just by repeating actions that lead to the reward of the sweet taste. It can learn to use entirely new and unnatural actions to obtain this reward and will be much better at exploiting a new source of food than would an animal limited to a fixed set of responses. As a practical example, cows will learn to press a button to turn on electrically driven rotating brushes, which they then stand next to, passively letting the brushes groom them (see the photo). The setup is highly unnatural, and the cows do not do the natural behavior of actively rubbing against trees or fences, but they clearly find being passively groomed by brushes highly rewarding because they will push open very heavy gates to get to them, repeatedly switch on the motors, and choose brushes over trees if given the choice (*12*). What matters to the cows is not doing the rubbing behavior but the usual result of doing the rubbing: the "reward" of the associated touch.

Understanding what animals find rewarding or aversive allows welfare to be defined by the animals themselves rather than having it imposed by humans. States of positive welfare (pleasure and contentment) can be objectively defined by what animals demonstrate that they find rewarding and will work for, such as by pressing buttons, pushing levers, or operating computer touch screens (technology now provides many different ways of finding out what animals value). They can demonstrate what stimuli they want and what behavior they value being able to perform, which may or may not include their natural behavior. It is even possible to ask animals how much they value being able to do different behaviors by comparing how hard they will work for them (for example, how many lever presses they will make, or what weights they will push) (*6*).

States of negative welfare can also be objectively defined as those produced by being forced to exist in conditions they find punishing (discomfort or pain) or in which they anticipate pain (fear) or do not obtain an expected reward (frustration). Differences between animals can be accommodated by separately investigating reward value of different outcomes to animals of different ages, breeds, and sexes, thus taking into account individual welfare needs. For example, adult laying hens will choose high perches for roosting at night (*13*), growing broiler chickens with their heavier bodies prefer lower more solid platforms (*14*), whereas very young chicks are brooded by their mother (if present) on the ground, not on perches at all. Simply providing the same sort of perches to all breeds and ages of birds and then judging their welfare by whether or not they perch is therefore to miss the animal's own point of view and what these different groups need and value for their welfare.

Judgements about animal welfare in different farming systems could easily include whether they provide what animals themselves value, according to what research shows they find rewarding or punishing. Knowledge of natural behavior is thus still important because it provides the most likely candidates for what animals might value, but we can now do much better than constructing farm environments only on the basis of whether animals are seen to be doing a specific set of species-typical behaviors. We have the means to look beyond what they do and to understand what they have evolved (or been selected by us) to regard as desirable environments. Understanding the nature of an animal means understanding not only what it does but what it values. Even John Webster, one of the originators of the Five Freedoms, came later to believe that the key Freedom should have been "Freedom to choose" (*15*).

Of course, animals do not always choose what is good for their health, so as with humans, what animals value in the short term has to be balanced against what is best for long-term health (*6*). We need on-farm research to find the optimum balance between these two important components of welfare.

CONCLUSIONS

Natural behavior as an animal welfare indicator gained popularity at a time when relatively little was known about what animals value. We now have much more information about what animals really want and have developed a wide range of techniques for letting them tell us. There have also been major advances in our knowledge of how the behavior of animals is motivated, how it develops, and how it differs between individuals. We are therefore in a good position to carry out research to test the validity of any claim that animal welfare has been improved by seeing whether the animals concerned have been provided with something they demonstrably value. The current tendency of the food industry to use natural behavior uncritically as an indication of good welfare without providing such evidence therefore needs to be challenged. Specifically, claims that animal welfare standards are being improved because animals are provided with opportunities for natural behavior such as access to pecking blocks requires, first, basic evidence that birds actually interact with and peck at these objects. Second and most important, there needs to be evidence that they value the opportunity to engage in the behavior sufficiently to be willing to work to be able to do it—that they will not only choose to do it but overcome obstacles or learn to perform unusual tasks to do so. The evidence needed is a combination of small-scale controlled studies on what animals want followed up by large-scale on-farm research to ensure that results are directly relevant to conditions experienced by animals on commercial farms and take account of variation between animals of different breeds and ages. Some of this evidence is already available but not used. Some will need a collaborative program of research between food producers and the research community. Either way, a more animal-centered view of welfare, replacing overreliance on the misleading concept of natural behavior, could result in genuinely improved standards of farm animal welfare. ■

REFERENCES AND NOTES

1. Food and Agriculture Organization of the United Nations (FAO), Statistics (2022); https://www.fao.org/statistics/en.
2. Farm Animal Welfare Council (FAWC), "Farm animal welfare in Great Britain: Past, present and future" (FAWC, 2009); https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/319292/Farm_Animal_Welfare_in_Great_Britain_-_Past__Present_and_Future.pdf.
3. D. J. Mellor *et al*., *Animals (Basel)* **10**, 1870 (2020).
4. Perdue, "Perdue commitments to animal care: Creating a culture of animal care" (2022); https://corporate.perdue-farms.com/responsibility/animal-care/commitments.
5. M. B. M. Bracke, H. Hopster, *J. Agric. Environ. Ethics* **19**, 77 (2006).
6. M. S. Dawkins, *The Science of Animal Welfare: Understanding What Animals Want* (Oxford Univ. Press, 2021).
7. F. Vanhonacker, W. Verbeke, E. Van Pouke, F. A. M. Tuyttens, *Livest. Sci.* **116**, 126 (2008).
8. W. H. Thorpe, "The assessment of pain and distress in animals, Appendix III. Report of the Technical Committee on Welfare of Animals in Intensive Livestock Systems" (HMSO, 1965).
9. McDonalds, Corporate statement on Animal Health and Welfare (2022); https://corporate.mcdonalds.com/corpmcd/our-purpose-and-impact/food-quality-and-sourcing/animal-health-and-welfare.html#ourPerformance.
10. G. J. Mason, *Trends Ecol. Evol.* **25**, 713 (2010).
11. E. T. Rolls, *Brain Computations and Connectivity* (Oxford Univ. Press, 2023).
12. E. McConnachie *et al*., *Biol. Lett.* **14**, 20180303 (2018).
13. S. Gunnarsson, J. Yngvesson, L. J. Keeling, B. Forkman, *Appl. Anim. Behav. Sci.* **6**, 7217 (2000).
14. M. Norring, E. Kaukonen, A. Valros, *Appl. Anim. Behav. Sci.* **184**, 91 (2016).
15. J. Webster, *Animals* **6**, 35 (2016).

ACKNOWLEDGMENTS

M.S.D. thanks K. Barger, K. Christensen, and E. Rolls for inspiring conversations. M.S.D. is principal investigator on a SMART BROILER grant from the Foundation for Food and Agricultural Research (FFAR), which supports research not reported here.

10.1126/science.ade5437



PERSPECTIVES

GEOCHEMISTRY

Earth's persistent thermostat

The dependence of rock weathering on temperature helps to steer Earth's climate

By **Robert G. Hilton**

For billions of years, Earth's surface climate has varied in temperature within a relatively narrow range, allowing liquid water to persist and ensuring the habitability of our rocky planet. The breakdown of rocks during weathering is thought to be central in keeping the planet warm enough, and not too cold, by providing a stabilizing feedback in the carbon cycle. On page 382 of this issue, Brantley *et al.* (*1*) reconcile measurements from the laboratory with those made across landscapes in soils and rivers to quantify the overall temperature dependence of this important feedback mechanism. They find that this process only works to stabilize temperatures when minerals are supplied quickly enough by erosion and if rainfall can quench the thirst of the weathering reactions. Thus, the role of weathering in steering Earth's climate is likely to have changed as the continents shifted and collided.

Carbon dioxide (CO_2) is a greenhouse gas that plays a major role in controlling the temperature at Earth's surface. Every year, volcanoes release ~0.1 petagrams of carbon (PgC) per year as CO_2 from the interior of the planet. Although this appears small compared with the ~44,000 PgC that the atmosphere and oceans contain, this volcanic leak of carbon quickly adds up over Earth's long history: In just a million years, volcanoes would release enough carbon to almost triple the atmosphere and ocean carbon stores. Left unchecked, this buildup of CO_2 would lead to runaway warming (*2*).

Instead, CO_2 is removed from the atmosphere when specific minerals in rocks are chemically broken down or weathered. The carbonic acid formed by mixing CO_2 with rainwater can dissolve silicate minerals. The weathering reactions remove CO_2 from the atmosphere and create bicarbonate, a dissolved form of carbon, which is flushed from soils into rivers and onward to the ocean. There, the weathering products are used to make new carbonate minerals that lock up the carbon. Altogether, the net impact of silicate weathering is the removal of CO_2 from the atmosphere. Prior research has revealed that the amount of CO_2 removal may be sensitive to temperature (*3*). If the surface of the planet is too warm, silicate weathering can remove more CO_2. If too cool, silicate weathering becomes more subdued and atmospheric CO_2 will build up. In this way, silicate weathering could act as Earth's thermostat (see the figure), keeping the climate hospitable for life to thrive (*2*, *3*).

The global temperature sensitivity of silicate weathering has remained difficult to quantify. This is because not only temperature matters. Minerals need to be supplied to be weathered, often by eroding fresh rocks and breaking apart new minerals (*4*, *5*). Water is important, too; it is needed for the reactions and to wash away the reaction products (*4*, *6*). A second challenge has been to reconcile laboratory-based measurements made over months (*7*) with chemical weathering rates measured in soils (*8*) and captured in the dissolved loads of rivers (*4*, *6*, *9*), which can record reactions happening over hundreds to thousands of years.

Brantley *et al.* used a thermodynamic framework to assess the temperature dependence of silicate weathering rates. Using calculations grounded in the Arrhenius equation that describes the relationship between reaction rate and temperature, they combined the numerous laboratory assessments with global soil profiles and river catchments. All of these datasets show an increase of silicate weathering with temperature, although the field rates increase more with temperature than those from experiments in the lab. This can be explained by an array of processes that happen in landscapes, including physical mechanisms such as fracturing, the formation of new minerals during weathering, and the role of microorganisms. Notably, Brantley *et al.* show that many locations around the world have a weak link between temperature and silicate weathering because either there is not enough mineral supplied to weather or they do not have enough water.

A convergent theme of the study of Brantley *et al.* and prior work (*4*, *6*, *9*) is that only if minerals are supplied fast enough can the silicate weathering thermostat work. Where erosion is low, the temperature sensitivity is muted or apparently nonexistent. Chemical weathering models and proxies of CO_2, climate, weathering, and erosion also

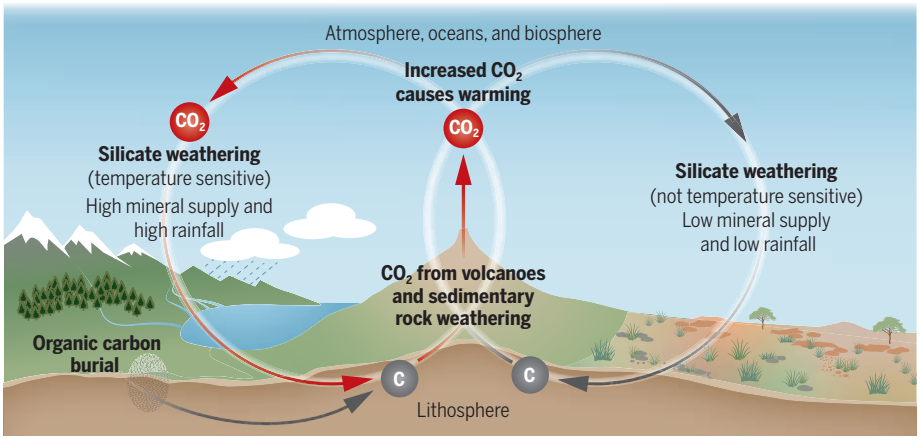
A view over Mafate circus in Réunion island is shown, where erosion and weathering of silicate rocks help feed into Earth's persistent thermostat.

PHOTO: AYMERIC BEIN/GETTY IMAGES

Department of Earth Sciences, University of Oxford, Oxford, UK. Email: robert.hilton@earth.ox.ac.uk

Silicate weathering thermostat

Over millions of years, atmospheric carbon dioxide (CO_2) is supplied by volcanic eruptions, as well as during sedimentary rock weathering. If CO_2 increases (red arrow), parts of Earth with high erosion and rainfall can increase CO_2 removal by silicate weathering. This provides a stabilizing feedback on climate over millions of years. However, silicate weathering fixes less CO_2 in the lithosphere if temperatures are cooler, mineral supply is limited, and rainfall is low.



provide support for these conclusions (*10*). For example, recent work has suggested that global erosion could have doubled over the past 10 million to 15 million years and made Earth's surface a more sensitive thermostat by supplying minerals and making reactions more sensitive to temperature and water supply (*10*), which affects climate and atmospheric CO_2.

The important role for erosion in the silicate weathering thermostat means that other factors complicate how it may steer Earth's climate. This is because erosion can supply other phases to the weathering zone (*11*) that release CO_2 when they break down. These phases include ancient organic matter in rocks and sulfide minerals that can be oxidized during weathering (*12*, *13*). These erosion and weathering-induced CO_2 emissions are globally important because they release as much CO_2 as volcanoes. In addition, recent work shows that the CO_2 release may increase with temperature (*13*). This seems to oppose CO_2 removal by silicate weathering. Furthermore, any holistic consideration of the long-term carbon cycle must also account for fluxes and drivers of CO_2 removal by organic matter burial, as has been suggested for the Himalayan erosional system (*14*). More research is needed to assess how the organic carbon cycle affects the thermostat settings.

Unfortunately, the natural silicate weathering process is too slow to help mop up the very large excesses of CO_2 that are released each year from human activities (~9.6 PgC $year^{-1}$). However, deliberately increasing silicate weathering by grinding up silicate minerals and applying them at a large scale to agricultural areas may provide a way to help reach net zero (*15*). It remains unclear whether the logistical challenges can be met for this enhanced weathering to contribute to CO_2 drawdown, but research into the fundamentals of mineral weathering reactions will be essential to make decisions and assess the water demands of sustaining the reactions (*1*).

The temperature dependence of silicate mineral weathering provides a persistent thermostat to guide Earth's long-term climate. However, changes in patterns of erosion, water supply, and organic carbon cycling have likely changed the thermostat's settings as the continents have drifted and collided over millions of years. Geochemical approaches in the laboratory, in the field, and in silico are crucial to further untangling the dynamics of the carbon cycle on Earth and other rocky planets. ■

REFERENCES AND NOTES

1. S. L. Brantley *et al.*, *Science* **379**, 382 (2023).
2. R. A. Berner, K. Caldeira, *Geology* **25**, 955 (1997).
3. J. C. G. Walker, P. B. Hays, J. F. Kasting, *J. Geophys. Res.* **86**, 9776 (1981).
4. A. J. West, A. Galy, M. Bickle, *Earth Planet. Sci. Lett.* **235**, 211 (2005).
5. M. E. Raymo, W. F. Ruddiman, P. N. Froelich, *Geology* **16**, 649 (1988).
6. K. Maher, C. P. Chamberlain, *Science* **343**, 1502 (2014).
7. A. F. White, S. L. Brantley, *Chem. Geol.* **202**, 479 (2003).
8. C. S. Riebe, J. W. Kirchner, R. C. Finkel, *Earth Planet. Sci. Lett.* **224**, 547 (2004).
9. B. Dupré *et al.*, *C. R. Geosci.* **335**, 1141 (2003).
10. J. K. Caves Rugenstein, D. E. Ibarra, F. von Blanckenburg, *Nature* **571**, 99 (2019).
11. R. G. Hilton, A. J. West, *Nat. Rev. Earth Environ.* **1**, 284 (2020).
12. M. A. Torres *et al.*, *Nature* **507**, 346 (2014).
13. G. Soulet *et al.*, *Nat. Geosci.* **14**, 665 (2021).
14. V. Galy *et al.*, *Nature* **450**, 407 (2007).
15. J. Hartmann *et al.*, *Rev. Geophys.* **51**, 113 (2013).

ACKNOWLEDGMENTS

R.G.H. acknowledges funding from the European Research Council Consolidator Grant "RIV-ESCAPE" (101002563).

10.1126/science.adf3379

MATERIALS CHEMISTRY

Decoding complex order in reticular frameworks

Periodicity and aperiodicity coexist in a nonchaotic, information-rich crystal structure

By **Omar M. Yaghi**[1,2] and **Zichao Rong**[1,2]

The synthesis of chemical structures has largely focused on achieving them as crystalline solids because of the relative ease with which crystals can be characterized at the atomic level using x-ray or electron diffraction techniques. Constituent atoms are arranged in a definite pattern of order that confers homogeneous (uniform) properties, which makes them attractive to industries ranging from energy to pharmaceuticals. What about amorphous solids, in which the atoms are arranged in an aperiodic fashion? Clearly, structures of this type are more complex and far more difficult to design and study because of their disorganized atomic arrangements. On page 357 of this issue, Meekel *et al.* (*1*) report an example of a crystalline material that combines periodicity with aperiodicity to create what they refer to as "complex order." This duality in structure can be leveraged to hold information that can be decoded, much like QR codes and barcodes.

Typically, a metal-organic framework (MOF) is an extended crystalline material composed of metallic ions or clusters that are building units linked by organic molecules. Meekel *et al.* generated an MOF with a metal unit that was rotationally disordered throughout the structure. The complex order that was generated is reminiscent of that created using Truchet tiles. Truchet tiles are not rotationally symmetric, and when placed squarely in a grid, they generate interconnecting patterns.

[1]Department of Chemistry, University of California, Berkeley, CA 94720, USA. [2]Bakar Institute of Digital Materials for the Planet, Division of Computing, Data Science, and Society, University of California, Berkeley, CA 94720, USA. Email: yaghi@berkeley.edu

GRAPHIC: N. CARY/*SCIENCE*

The complex order in the Truchet MOF of Meekel *et al.*, which they call TRUMOF-1, is realized by having a rotationally disordered metal zinc and organic linker (1,3-benzenedicarboxylate) that propagates across the crystal in an aperiodic fashion while still being affixed to an otherwise ordered lattice of oxygen (see the figure). What makes TRUMOF-1 distinctive is that the disordered and ordered constituents are covalently linked in a three-dimensional structure that is also porous. This MOF is a major addition to a larger family in which such complex order provides for anisotropy and directionality in crystals (*2*). In considering the amount of chemical information that periodic and aperiodic atomic arrangements contain, aperiodic structures are far richer and arguably more interesting because they exhibit more diversity, variation, and complexity. However, the lack of a means to characterize aperiodic materials has prevented the full exploitation of their potential as chemical information storage systems. The establishment of reticular chemistry—the linking of molecular building blocks with strong bonds to create extended crystalline structures—offers an entry point to addressing the characterization problem by making structures such as MOFs in which periodic and aperiodic atomic arrangements are combined, giving rise to emergent properties (*3*).

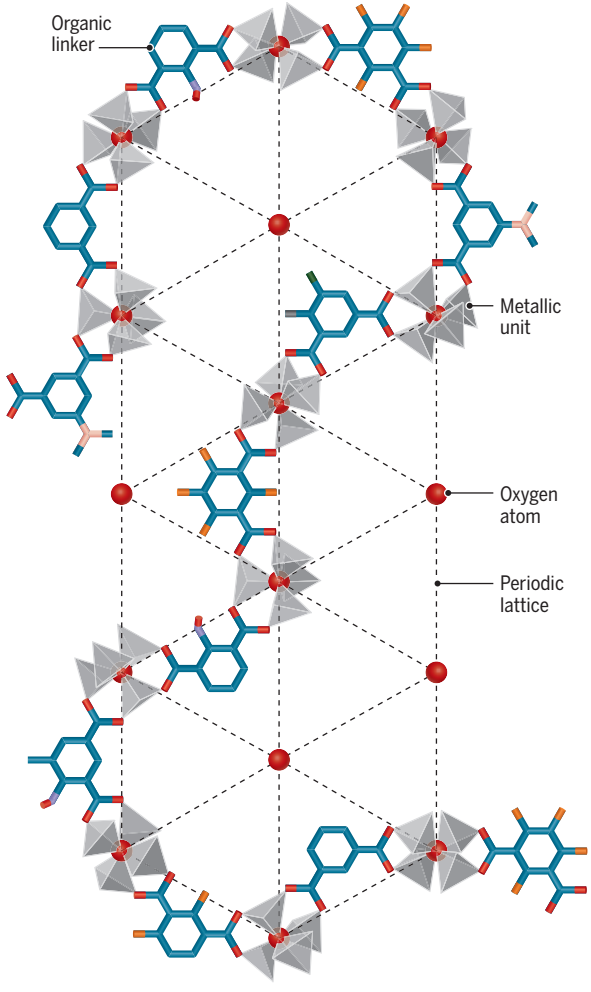
Given the flexibility with which reticular frameworks such as MOFs can be designed, variability of constituents, and ability to chemically modify their pores, complex order appears to be a natural outcome of this ultrahigh variability, especially when applied to a single framework. Thus, an MOF structure can be investigated at the atomic scale in detail to immediately articulate forms of complex order. This is a typical computational approach for characterizing a complex system when no physical techniques exist otherwise. There are several examples of structures in which periodicity and aperiodicity coexist to embody the desirable qualities of both. DNA is a perfect example of complex order in one dimension. The polyphosphate sugar chain serves as the repetitive and periodic backbone onto which aperiodic spatially arranged nucleotides are covalently bound. The complex order in DNA is based on variation in the composition of the rungs of the double helix as represented by nucleotide pairs, whereas in a Truchet MOF, it is based on the relative configurations of the constituent units.

There are other forms of complex order. In multivariable MOFs, up to 10 different metal ions are incorporated into their metallic units to generate a heterogeneous spatial arrangement of metals in the form of, for example, metal sequences that contain short or long metal repeats (*4*–*6*). When eight different organic moieties are covalently attached to a three-dimensional MOF, a "guest" molecule that passes through the pores of the MOF can sample a specific sequence of functional groups across the entire crystal (*7*). This can impart gas selectivity properties onto the MOF, which allows the whole to be greater than the sum of its parts—that is, the way that the constituents of the MOF combine adds a different quality to the final complex structure. The importance of complex order in catalysis is also revealed in a reticular structure that incorporates a heterogeneous arrangement of short peptides. Notably, these crystals carry out selective amide bond cleavage akin to an enzyme (*8*).

Although the complex order in the Truchet MOF of Meekel *et al.* was elegantly decoded by analyzing the diffuse x-ray scattering of the crystals, it is not always possible to apply this method to other MOFs because of their openness and large unit cells. Attempts to use atom probe tomography (*4*) and solid-state nuclear magnetic resonance (*9*) techniques have only facilitated decoding the sequence (that is, the spatial order) of constituent metal units and organic linkers down to the nanometer level, which leaves many questions unanswered about the design of properties in reticular structures that bear complex order. Decoding such complex order on a routine basis requires moving away from the traditional practice of making, characterizing, and measuring properties of a structure according to methods in which inputs into what is made are related to outputs. Rather, materials chemistry must decode the correlation between the inputs and outputs. This understanding should reveal the rich information that is held by structures of complex order. The execution of this vision can be achieved by blending the chemistry and synthesis of materials with computing and data science fields. This approach will most certainly require machine-learning tools (*10*). The Truchet MOF demonstrated by Meekel *et al.* highlights the power of using creative thinking to decode complex order. ■

Duality in structure

The building units of a metal-organic framework can be linked to form aperiodic, rotationally variable configurations. These organic linkers (which bear variability in functional groups) and metallic units can be superimposed on (and strongly bonded to) a periodic lattice.



GRAPHIC: YAGHI *ET AL.* ADAPTED BY K. HOLOSKI/*SCIENCE*

REFERENCES AND NOTES

1. E. G. Meekel *et al.*, *Science* **379**, 357 (2023).
2. W. Xu *et al.*, *Nat. Rev. Mater.* **5**, 764 (2020).
3. H. Furukawa, K. E. Cordova, M. O'Keeffe, O. M. Yaghi, *Science* **341**, 1230444 (2013).
4. Z. Ji, T. Li, O. M. Yaghi, *Science* **369**, 674 (2020).
5. L. J. Wang *et al.*, *Inorg. Chem.* **53**, 5881 (2014).
6. C. López-García *et al.*, *J. Am. Chem. Soc.* **144**, 16262 (2022).
7. H. Deng *et al.*, *Science* **327**, 846 (2010).
8. A. M. Fracaroli *et al.*, *J. Am. Chem. Soc.* **138**, 8352 (2016).
9. X. Kong *et al.*, *Science* **341**, 882 (2013).
10. D. Rolnick *et al.*, arXiv:1906.05433v2 (5 November 2019).

10.1126/science.adf0261

IMMUNOLOGY

A vaccine sanctuary in the lymph node

B cell follicles in the lymph node protect vaccines to enhance immune responses

By **Anurag Kalia** *and* **Heather D. Hickman**

After injection, vaccines drain to the lymph node, where naïve B cells in lymph node follicles recognize vaccine antigens, become activated, and mutate immunoglobulin genes to create protective antibodies. The remarkable success of vaccines suggests that this process occurs efficiently and results in antibodies against vaccine antigens that have retained their native structure. However, during the journey from injection site, vaccines must run a gauntlet of extracellular proteases before reaching follicular B cells. It is unclear how antigens are preserved as they are shepherded into B cell follicles and, after arrival, how they are maintained in a structurally intact conformation for continued B cell access. On page 350 of this issue, Aung *et al.* (*1*) reveal that lymph node B cell follicles in mice possess low intrinsic protease activity, allowing follicular dendritic cells (FDCs) to retain vaccine in its native form and potentially revealing ways to enhance vaccine efficacy.

After deposition into the tissue, protein-based vaccines enter lymphatic vessels, which propel fluid, particles, and cells from peripheral sites to the closest lymph node. Lymph-borne vaccine particles first enter the lymph node in a distinct microenvironment called the subcapsular sinus (SCS) (see the figure). The SCS is filled with lymphatic fluid and lined with a specialized subset of macrophages (SCS macrophages) that can capture lymph-borne immune complexes and pass them on to underlying B cell follicles (*2*). Despite a capacity to transport antigens, SCS macrophages also internalize particulate antigens (including lymph-borne viruses) and express intracellular and secreted proteolytic enzymes (*2, 3*). Additionally, protease activity has been detected in the SCS after injection of substrates that fluoresce owing to protease cleavage (*4*).

Laboratory of Clinical Immunology and Microbiology, National Institutes of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA.
Email: hhickman@mail.nih.gov

Recent advances in microscopy have enabled the precise spatiotemporal mapping of vaccines and other antigens in the lymph node. Although SCS macrophages access fluorescently labeled nanoparticle vaccines displaying glycosylated viral proteins, vaccines are cleared from most nodal microanatomical niches, including the SCS, within a few days of injection. One notable contrast to widespread nodal vaccine loss over time is retention on FDCs (*5*). Unlike conventional dendritic cells, FDCs are nonhematopoietic stromal cells that are located in B cell follicles. FDCs bind and trap immune complexes that are relayed by SCS macrophages (*6*), retaining antigen in its native structure. FDCs can also endocytose and recycle antigens, which is thought to preserve antigen from degradation by extracellular proteases (*7*). Some viruses, most notably HIV, can also access the FDC network and use this protective capacity to establish a long-term source of replication-competent virus (*8*). Both sugars and immune complexes can direct antigen to FDCs (*9*), which then support the formation of germinal centers (sites for antibody somatic hypermutation within B cells in the follicle) (*10*).

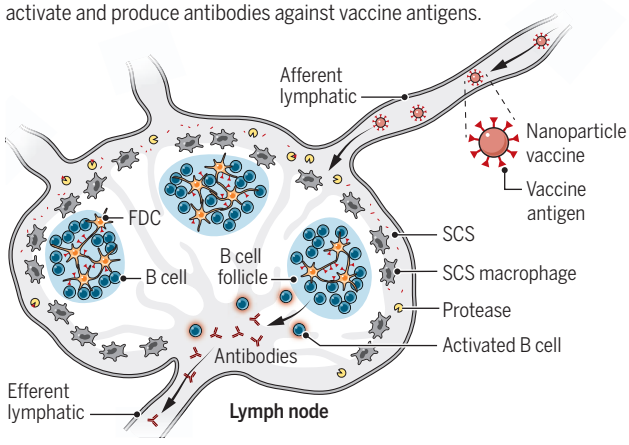
Building on their previous studies demonstrating that mannose-binding lectin on SCS macrophages and complement direct viral protein–displaying nanoparticles to FDCs (*5*), the authors questioned how FDCs were able to maintain vaccine antigen when cells in other areas of the node could not. They hypothesized that extracellular proteases might degrade antigen unless it was protected by either a cell or the microenvironment. After immunization of mice, they found that fluorescently labeled nanoparticles first accumulated in the SCS and remained intact, as expected. However, over time, spatial patterns of nodal nanoparticle degradation were revealed, with those in the SCS or interfollicular regions being degraded, whereas nanoparticles associated with FDCs were not. These results confirmed previous findings of FDCs sheltering native antigen, but did not discern whether this was due to a specific FDC property or a function of the environment.

In an innovative set of experiments, Aung *et al.* queried the protective capacity of the B cell follicle itself. They first directed antigen into either the B cell follicle or the T cell zone of the lymph node by antibody-mediated targeting. Whereas antigen present in the T cell zone was rapidly degraded, follicular antigen was preserved—even when associated with B cells rather than FDCs. When the authors artificially increased protease levels in the follicles, FDC preservation of intact antigen was substantially reduced. Thus, although FDCs may recycle antigen to prevent degradation, the follicles in general also contribute to native antigen maintenance by forming a low-protease refuge within the node.

The authors next applied this knowledge to design and test the specificity of humoral immune responses when antigen was targeted to the B cell follicles to avoid degradation in the SCS. Immunization strategies that directed antigen to FDCs resulted in large increases in the number of germinal center B cells specific for intact antigen. The authors achieved this by using nanoparticles that are rapidly transported to FDCs through the mannose-binding lectin-complement pathway. This corresponded to increased levels of antibodies that provide vaccine-mediated protection when compared to non-FDC–targeted vaccines. Antibodies specific for de-

Preventing vaccine antigen degradation

Nanoparticle vaccines encounter extracellular proteases en route to B cell follicles in the lymph node. Although vaccine antigen is rapidly degraded in the subcapsular sinus (SCS) region of the lymph node, B cell follicles have low protease activity. This preserves antigen integrity and allows follicular dendritic cells (FDCs) to present intact antigen to B cells, which activate and produce antibodies against vaccine antigens.



GRAPHIC: K. HOLOSKI/*SCIENCE*

graded antigen did not increase even with enhanced follicular vaccine delivery, demonstrating that the observed effect was not simply a result of increased follicular antigen levels. Targeting the mannose-binding lectin-complement pathway could represent a general strategy for protein-based vaccination to enhance protective immune responses to the most appropriate antigen.

These discoveries have exciting translational implications and reveal FDC function; however, several questions remain. From an evolutionary perspective, what is the advantage of high protease activity in nonfollicular areas of the lymph node? Degradation of viruses in the SCS could serve as a mechanism by which the lymph node prevents infection. However, in mouse models, SCS macrophages are infected by many viruses (*3*), suggesting that if proteolytic pathogen inactivation occurs, it is generally not sufficient to eliminate infection. Furthermore, once infection is established, the lack of protease activity may enable follicles to serve as refuges for viruses. In situations where FDCs trap and retain virus, strategies providing short-lived, targeted enhancement of follicular protease activity might represent a therapeutic approach to purge infection.

Historically, successful vaccines required little or no knowledge of antigen movement and retention. Recent efforts, including those by Aung *et al.*, have generated a deeper understanding of the impact of antigen preservation on the development of protective antibodies against structurally relevant antigens. Future studies will determine whether these same principles of follicular targeting are advantageous during human vaccination. Regardless, a roadmap of safe and hazardous zones for antigen integrity within the mouse lymph node stands to inform new strategies for vaccine design and enhancement. ■

REFERENCES AND NOTES

1. A. Aung *et al.*, *Science* **379**, eabn8934 (2023).
2. T. G. Phan, J. A. Green, E. E. Gray, Y. Xu, J. G. Cyster, *Nat. Immunol.* **10**, 786 (2009).
3. A. Bellomo, R. Gentek, M. Bajénoff, M. Baratin, *Cell. Immunol.* **330**, 168 (2018).
4. D. M. Catron, K. A. Pape, B. T. Fife, N. van Rooijen, M. K. Jenkins, *J. Immunol.* **184**, 3609 (2010).
5. T. Tokatlian *et al.*, *Science* **363**, 649 (2019).
6. J. G. Tew, T. E. Mandel, A. W. Burgess, *Cell. Immunol.* **45**, 207 (1979).
7. B. A. Heesters *et al.*, *Immunity* **38**, 1164 (2013).
8. B. A. Smith *et al.*, *J. Immunol.* **166**, 690 (2001).
9. R. Roozendaal, M. C. Carroll, *Immunol. Rev.* **219**, 157 (2007).
10. X. Wang *et al.*, *J. Exp. Med.* **208**, 2497 (2011).

ACKNOWLEDGMENTS

The authors are supported by the National Institute of Allergy and Infectious Disease's Division of Intramural Research.

10.1126/science.adf5134

HYPOTHESIS

Infection infidelities drive innate immunity

Unsuccessful pathogenic activities may be key to activating host defenses to infection

By **Jonathan C. Kagan**

Within multicellular organisms, innate immune pattern recognition receptors (PRRs) control host defenses to infection. These receptors are activated by microbes, which facilitate their own demise by producing PRR ligands called pathogen-associated molecular patterns (PAMPs). Notably, the mechanisms of PRR-mediated microbial detection are inconsistent with the sensing of successful pathogens. I propose that PRRs do not detect pathogenic agents per se. Rather, PRRs detect PAMPs that are released from microbes as a result of biochemical infidelities, or mistakes, that occur during infection. These mistakes render individuals within an otherwise infectious population noninfectious but immunostimulatory. Microbes could evolve strategies that increase the fidelity of their infectious strategies to evade PRRs. However, imperfect activities enable biochemical innovations that ensure the survival of the species. By detecting PAMPs that are released as a result of low-fidelity biochemical activities, PRRs may, in effect, be targeting the very process that microbes need for long-term survival—evolvability.

> **"...the mechanisms of pattern recognition receptor–mediated microbial detection are inconsistent with the sensing of successful pathogens."**

Examples of PRRs include Toll-like receptors (TLRs), RIG-I–like receptors (RLRs), the enzymes cyclic guanosine monophosphate–adenosine monophosphate synthase (cGAS) and protein kinase R (PKR), and select members of the nucleotide-binding and leucine-rich repeat containing (NLR) protein family. The most intuitive examples of PRR-mediated detection of infection infidelities derive from studies of nucleic acid–sensing TLRs (*1*). These TLRs reside in the endosomal network, with their ligand-binding domains oriented into the endolysosomal lumen. This subcellular positioning is not conducive to the detection of viable microbes because lysosomes are a subcellular site of microbial destruction (see the figure).

The specific nucleic acids detected by TLRs suggest preferential detection of dead microbes. Free guanosines and uridines represent cofactors that, along with single-stranded RNA (ssRNA), activate TLR7 and TLR8, respectively (*1*). Similarly, trinucleotides containing cytosines synergize with CpG-containing ssDNA to activate TLR9. The requirement of short nucleotide sequences (or free nucleosides) to activate several TLRs supports the idea that lysed microbial cells or degraded virions are detected after their genomes have been hydrolyzed. Even when free nucleosides are not required for TLR activation, as in the case of TLR3, the nature of the ligand is not conducive to the detection of viable microbes. Crystal structures of TLR3 revealed extensive contacts with double-stranded RNA (dsRNA) (*2*). The genomes of dsRNA viruses are typically enclosed in chromatin-like ribonucleoprotein (RNP) complexes, which should prevent interactions between the genome and TLR3.

Endosomal TLRs that specifically detect bacterial nucleic acids also do not detect viable microbes. For example, mouse TLR13 and human TLR8 recognize a specific RNA sequence that is only present in bacterial 23*S* ribosomal RNA. Only after bacterial degradation in lysosomes would TLR13 or TLR8 be exposed to their RNA ligands. Plants also detect a component of the bacterial translation machinery, the highly conserved elongation factor thermo unstable (EF-Tu) protein. The plant PRR EF-Tu receptor (EFR) preferentially detects proteolytic fragments of EF-Tu that are in the extracellular space as opposed to the intact EF-Tu protein (*3*). Lysed or stressed

Harvard Medical School and Division of Gastroenterology, Boston Children's Hospital, Boston, MA, USA.
Email: jonathan.kagan@childrens.harvard.edu

bacterial cells are a likely source of extracellular EF-Tu that may be subsequently hydrolyzed to create optimal PRR ligands. Bacteria that executed successful virulence strategies would have avoided destruction by mammalian lysosomes or plant defenses, thereby maintaining their intracellular pools of nucleic acids and EF-Tu and avoiding PRR detection.

Cytosolic nucleic acid–sensing PRRs also operate in ways that are not conducive to detection of successful pathogens. For example, cGAS, absent in melanoma 2 (AIM2), and interferon-γ inducible protein 16 (IFI16) are most efficiently activated by non–protein-coated dsDNA (*4*), and PKR inactivates mRNA translation upon detection of noncoated RNA ligands as opposed to RNPs (*5*). In the case of cGAS, this PRR can interact with protein-coated DNA, but the downstream interferon-inducing activities of this enzyme only occur when cGAS binds protein-free DNA (*6*). The chromatin-like nature of replicating viral DNA should prevent DNA-sensing PRRs from detecting successful pathogens. PRR activation during DNA virus infection may therefore result from the detection of noncoated (and perhaps inactivated) virions.

The ability of the RLR RIG-I to detect infection is dependent on its recognition of cytoplasmic di- or triphosphorylated RNAs (*5*). Yet, several viral pathogens modify their RNA termini (through capping or dephosphorylation) in a manner that is not compatible with RIG-I detection (*7*). MDA5, the other immunostimulatory RLR, does not recognize the termini of RNA and rather binds internal structures of dsRNA that are typically encapsulated in RNPs (*5*). In several cases, immunosuppressive but productive viral infections can be rendered nonproductive (and immunostimulatory) when expression of the histone-like component of the RNP is experimentally restricted (*8*). Thus, successful pathogens are able to prevent exposure of their PAMPs to PRRs or block downstream PRR signaling through virulence factor activities.

Despite virulence strategies that encapsulate viral genomes into RNPs and modify the 5′ termini of their nucleic acids, RNA viruses typically stimulate RLRs (*5*). This disconnect between pathogenesis strategies that should prevent PRR signaling and the known ability to stimulate PRRs can be explained by the selective detection of the errors made at some stage of the infectious cycle. Low-fidelity RNA polymerases or inefficiencies of capsid assembly may disrupt the life cycle in subsets of the infectious inoculum, leading to PRR activation. Consistent with this idea, deep sequencing studies of infected cells and animals revealed preferential interactions between RIG-I and short influenza virus RNAs that do not encode functional proteins, as opposed to full-length viral genomes (*9*).

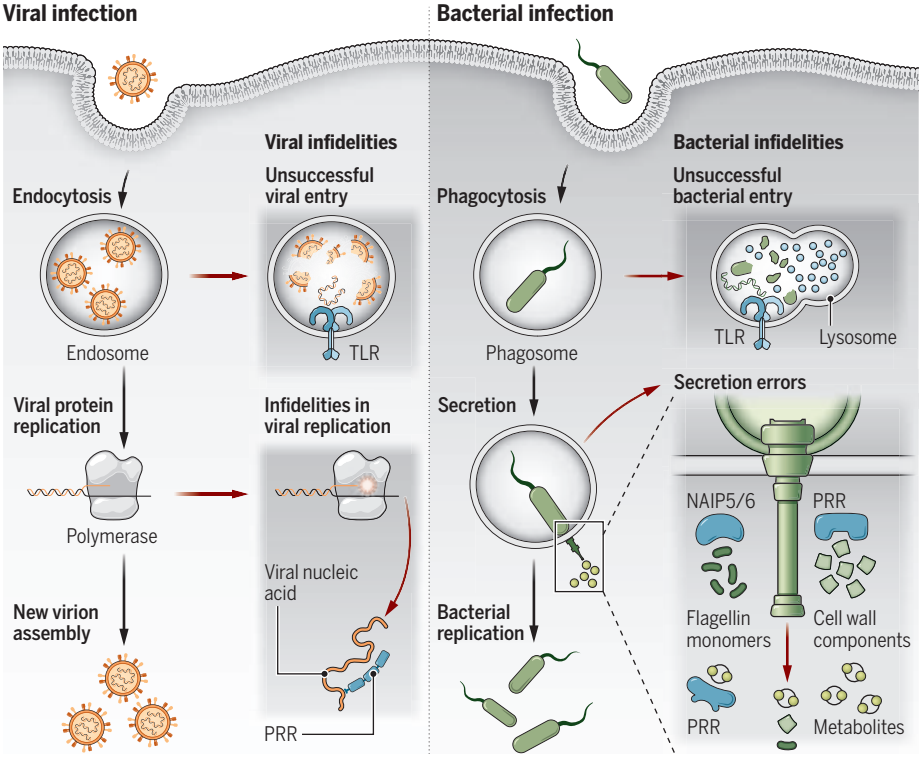
The bacterial flagellar apparatus, the type III secretion system (T3SS), and T4SS are each multiprotein complexes that span the inner and outer bacterial membranes (*10*). These complexes export bacterial proteins into the extracellular space. In the case of the flagellar system, the protein secreted is flagellin, which subsequently oligomerizes into a flexible filament (the flagellum) that enables bacterial motility. Flagellin is a TLR5 ligand, but the TLR5 binding site is positioned at the interface between adjacent flagellin monomers (*11*). A functional flagellum will therefore not display its TLR5 binding site to this PRR. Yet, flagellin monomers that fail to incorporate stably into the flagellum can activate TLR5. Similar innate immune strategies exist in plants, where the PRR flagellin-sensitive 2 (FLS2) binds proteolytic fragments of flagellin monomers at sites that are not exposed on the surface of an intact flagellum (*12*). Biochemical inefficiencies in flagellar biosynthesis or function (or an inability to evade cleavage by extracellular plant proteases) likely explain how these PRRs access their flagellin ligands.

T3SS and T4SS complexes inject bacterial proteins into eukaryotic cells, resulting in host manipulation during infection (*10*). These systems are thought to also inject flagellin (*13*). Assuming that the only function of flagellin is to build a motility organelle, then its injection into a host cell represents an error in secretion system substrate selection. Upon injection into the cytosol, the PRRs NLR family apoptosis inhibitory protein 5 (NAIP5) and NAIP6 bind to flagellin monomers. In a thematically similar manner, NAIP1 and NAIP2 bind to the monomeric building blocks of the needle and inner rod components of T3SS from several bacteria (*13*). Similar to flagellin monomers, these injected monomers can be considered useless for secretion system function.

Bacterial cell wall lipopolysaccharide (LPS) and peptidoglycan components, such as muramyl dipeptide (MDP), can also be injected into host cells through secretion system activities, as can adenosine diphosphate (ADP)–heptose, a precursor of LPS

Activating pattern recognition receptors

Successful viral and bacterial infections do not activate pattern recognition receptors (PRRs) and downstream innate immune responses. Instead, infidelities in viral and bacterial infections can create ligands for PRR activation. For example, nucleic acids from pathogen degradation in endosomes and lysosomes can activate Toll-like receptors (TLRs). Mistakes in viral replication can also activate PRRs. Moreover, bacterial secretion systems can mistakenly release cell wall components, metabolites, or flagellin monomers into the cytosol, which stimulate PRRs, such as NLR family apoptosis inhibitory protein 5 (NAIP5).



GRAPHIC: A. FISHER/*SCIENCE*

(*14*). Each of these cell wall components stimulate distinct PRRs. It stands to reason that the injection of these biosynthetic intermediates into the host cell does not serve the purpose of building a bacterial cell wall. Additionally, in plants, medium-chain 3-hydroxy fatty acids that serve as LPS biosynthesis precursors can activate immune responses through the *Arabidopsis thaliana* PRR lipooligosaccharide-specific reduced elicitation (LORE) (*15*). PRRs that detect cell wall biosynthesis intermediates may therefore be operating to detect biochemical errors in cell wall biosynthesis or stability. A similar theme is offered by bacterial cyclic dinucleotides, which serve a metabolic function in bacteria yet activate the PRR stimulator of interferon genes (STING) upon their release into mammalian cells (*4*).

Collectively, >24 PRRs in plants and animals are preferentially activated by PAMPs that are released through low-fidelity infection strategies. These strategies, if executed properly, should prevent microbial degradation in lysosomes, the inefficient encapsulation of viral nucleic acids, and the mistaken injection of cell wall components, along with virulence factors, into the host cell cytosol. The result of a successful infection is therefore the evasion of PRR detection. Increasing infection fidelity could maximize PRR evasiveness in the short term but may limit adaptability of the species, with implications for host diversification.

Like all areas of biology, exceptions to suggested biological rules exist. For example, PRRs that detect bacterial LPS molecules and lipoproteins may represent examples of pathogen detection systems. These molecules are displayed on the surface of bacteria and are detected by specific TLRs in their intact forms. Similarly, the PRRs dectin-1 and dectin-2 recognize β-glucans and α-mannans, respectively, whose positioning on the fungal cell wall facilitates detection. However, there are possible explanations for these exceptions. The mammalian LPS detection systems are not as ancient as other PRR networks, such as the endosomal TLRs, RLRs, cGAS, and STING. These ancient PRR systems detect the most definitive example of PAMPs that are released as a result of infection infidelity—nucleic acids. Outside of mammals, few examples of immunostimulatory LPS activities (or LPS receptors) exist. Indeed, the genes encoding the mammalian LPS receptors CD14, TLR4, and MD2 are absent from the genomes of fish and invertebrates (which constitute >90% of metazoan life). Thus, exceptions to the error-centric theme of pattern recognition may be a recent evolutionary innovation.

For the targeting of infection infidelities to be a successful defense strategy, these infidelities must be common. The widespread presence of viral particles containing incomplete genomes during infection and the injection of cell wall components into mammalian cells by diverse bacterial secretion systems supports this contention. Additionally, the preference of plant PRRs to detect protein and peptidoglycan degradation products as PAMPs suggests that low-fidelity metabolic or virulence activities are prevalent and can drive innate immunity.

Even if low-fidelity biochemical reactions were rarely detected by PRRs during infection, the rarity would need to approach zero for host defenses to be avoided. The very nature of the mammalian inflammatory and plant hypersensitive responses induced by PRRs enables tissue (and likely systemic) defenses to be executed against the entire microbial community. Studies of infection at the single-cell level could reveal the generality of these concepts and how they affect defense. Additional work is also needed to understand how inflammatory activities that drive tissue repair after infection resolution may be affected by residual PAMPs derived from dead microbes.

An implication of the infection infidelity concept is that it provides the opportunity to develop anti-infective drugs that are also immunotherapies. It is possible that drugs that disrupt viral life cycles would, in immunocompetent cells, result in the release of PAMPs to the host immune system. Similarly, antibiotics that target bacterial cell walls may cause the release of cell wall components or cyclic dinucleotides that stimulate PRRs. Integrating PRR assays in anti-infective clinical pipelines may therefore offer opportunities for drug development. ■

REFERENCES AND NOTES

1. N. A. Lind *et al.*, *Nat. Rev. Immunol.* **22**, 224 (2022).
2. J. Choe *et al.*, *Science* **309**, 581 (2005).
3. G. Kunze *et al.*, *Plant Cell* **16**, 3496 (2004).
4. X. Tan *et al.*, *Annu. Rev. Microbiol.* **72**, 447 (2018).
5. S. Hur, *Annu. Rev. Immunol.* **37**, 349 (2019).
6. C. Zierhut *et al.*, *Cell* **178**, 302 (2019).
7. E. Decroly *et al.*, *Nat. Rev. Microbiol.* **10**, 51 (2012).
8. B. E. Nilsson-Payant *et al.*, *J. Virol.* **95**, e02274-20 (2021).
9. A. J. W. te Velthuis *et al.*, *Nat. Microbiol.* **3**, 1234 (2018).
10. J. E. Galán, G. Waksman, *Cell* **172**, 1306 (2018).
11. K. D. Smith *et al.*, *Nat. Immunol.* **4**, 1247 (2003).
12. Z. Bauer *et al.*, *J. Biol. Chem.* **276**, 45669 (2001).
13. J. von Moltke *et al.*, *Annu. Rev. Immunol.* **31**, 73 (2013).
14. P. Zhou *et al.*, *Nature* **561**, 122 (2018).
15. A. Kutschera *et al.*, *Science* **364**, 178 (2019).

ACKNOWLEDGMENTS

Thanks to members of the Kagan laboratory and V. Hornung for helpful discussions. J.C.K. is supported by National Institutes of Health grants AI167993, AI116550, and P30DK34854. J.C.K. holds equity in Corner Therapeutics, Larkspur Biosciences, and Neumora Therapeutics.

10.1126/science.ade9733

NEUROSCIENCE

Neurons share an intense load

Time-of-day–detecting cells in the eye give customized responses to light intensity

By **Paul R. Martin**

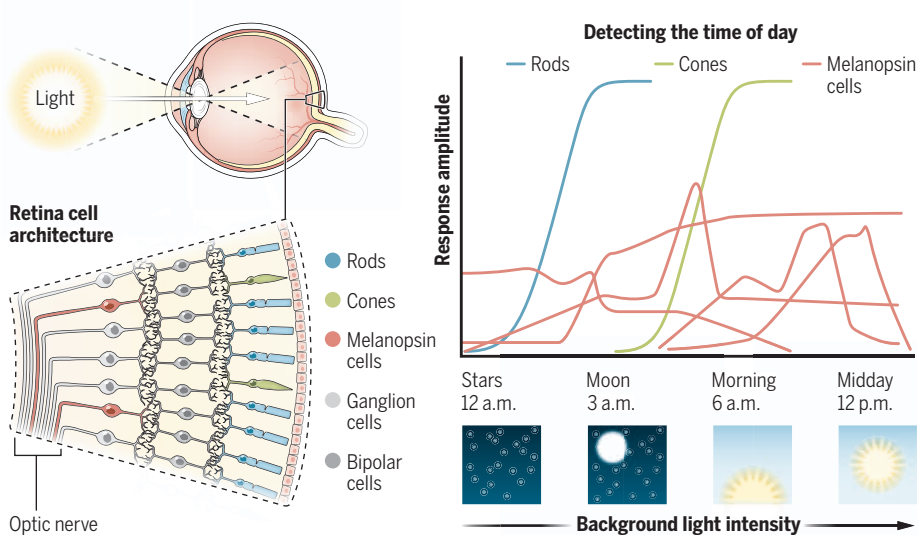
Tiny pulses of electrochemical energy known as "spikes" underlie brain function, from sensation and cognition to engendering vigorous action or quiet reflection. But exactly what kind of messages do spikes transmit to, through, and from the brain? On page 376 of this issue, Liu *et al.* (*1*) show how spiking activity in a small set of neurons in the macaque monkey eye can inform the brain about the huge range of environmental illumination encountered across every 24-hour day-night cycle. Light detection by the eye can synchronize the body's biological (or circadian) rhythms to this cycle, regulating essential functions such as sleep, attention, and energy expenditure. The authors found that unlike conventional photoreceptors (rods and cones, which show stereotyped responses across a limited range of background intensities), each time-of-day–detecting neuron signals a different range of photon flux, so that together the population can encode the vast intensity range from starlight to bright sunlight.

A central problem addressed by Liu *et al.* relates to a long-established observation: An individual neuron has limited signaling bandwidth. About a century ago, pioneering neuroscientists Lucas and Adrian (*2*, *3*) first found evidence that neurons transmit information in the form of electrochemical impulses. Adrian later followed Lucas's suggestion (*4*) to build a thermionic valve amplifier and measure these impulses in single stretch-receptor fibers from frog sartorius muscle. He found that when the muscle was stretched, the rate of impulses increased in proportion to the degree of stretch. The ensuing rate-coding hypothesis has become a central tenet in understanding the messages carried by individual neurons. For example, high spike rates in sensory neurons engender intense sensations, and if spikes are widely spaced

The University of Sydney Save Sight Institute, Sydney, Australia. Email: paul.martiin@sydney.edu.au

Environmental radiance coding in primate retina

The retina comprises photoreceptors (rods and cones) that show stereotyped responses across a limited range of background light intensities. In addition, melanopsin-expressing retinal output neurons (retinal ganglion cells called melanopsin cells) show more complex and idiosyncratic response functions with peak amplitudes of receptor responses according to different environmental irradiances.



in time, the sensations are correspondingly feeble. Likewise, high spike rates in neurons that innervate muscles generate strong contractions and low spike rates generate only weak contractions.

But the rate coding is constrained by biology: Each spike uses energy, neurons typically do not sustain rates above 50 spikes per second, and the number of discriminable levels of activity observed in a single neuron is correspondingly low (*5*). Spike-rate adaptation (where spike rate slowly falls if stimulus intensity does not change) can broaden nerve signal bandwidth (*5*) but comes at the cost of favoring detection of rapid changes over slow changes in the environment. How, then, can the brain be informed of the very slow (24-hour period) and very large (over eight logarithmic units) changes in environmental light intensity encountered by most animals across a day-night cycle?

Nerve signals serving the sense of sight are carried along the optic nerve to the brain on the extended processes of retinal output cells called ganglion cells. An important clue to understanding time-of-day coding came when a new ganglion cell type was identified in 2002 (*6*, *7*). These ganglion cells, known as melanopsin cells, show intrinsic photosensitivity based on a melanopsin-initiated phototransduction cascade, and innervate brain centers that control circadian rhythms (*6*–*10*). Melanopsin belongs to a large family of light-sensitive opsins. In the eye, the best known are in rod and cone photoreceptors. By contrast, the melanopsin cells have many features in common with the rhabdomeric photoreceptors expressed in the eyes of invertebrates. To study the melanopsin cascade, Liu *et al.* developed an antibody-based method to identify melanopsin cells in live ex vivo macaque monkey retina, as well as in mouse cells carrying transgenically inserted components of the human melanopsin cascade. They found that the monkey melanopsin cells have idiosyncratic irradiance response functions, with distinct cells showing activity peaks at different levels of absolute photon flux (called population-radiance encoding). Thus, the population of melanopsin cells work together to signal a larger range of intensity than can be covered by a single neuron rate code (see the figure).

Axonal recordings were previously used to show a similar population-coding mechanism in mouse retina (*11*). But the live immunotagging method developed by Liu *et al.* goes much further, allowing observations of phenomena such as multistable photoswitching in melanopsin cells. When rod and cone opsins absorb a photon, they fall into an inactive state, and the chromophore (light-absorbing molecule) 11-*cis*-retinal must pass through a multistep reactivation cascade involving transport out of, then back into, the photoreceptor. By contrast, the chromophore in melanopsin cells can be reversibly knocked into or out of active states by successive photon absorption. This and other features of melanopsin-based phototransduction extend the range of light intensity encoded by individual melanopsin cells.

In addition to revealing population coding in the nervous system, the results of Liu *et al.* open a door to manipulating melanopsin cell activity by tailoring the intensity and spectral content (roughly, the brightness and color) of artificial lighting, including therapeutic lighting for mood disorders such as seasonal affective disorder (*12*), and for reducing disturbances to circadian rhythms experienced by shift workers, new parents, and in air travel across time zones. Activated melanopsin induces persistent spiking in melanopsin cells, which Liu *et al.* show can be preferentially driven back to an inactive (nonspiking) state by long-wavelength light. This observation points to the potential for long-wavelength light to (paradoxically) mimic the effects of low light levels at nighttime by shutting down melanopsin cells.

It is important to emphasize that the study by Liu *et al.* goes beyond simply reproducing in monkeys what has been previously found in mice. Mice and monkeys occupy distinct ecological and behavioral niches and, crucially, mice are primarily nocturnal whereas monkeys (and most humans) are primarily active during the day. This means that a common proximate mechanism (population-radiance encoding) drives distinct circadian behaviors, sending mice to hide and sleep at illumination levels where monkeys start to rise and begin their daily activity. Every physiological process in animal bodies undergoes circadian rhythms (*13*). The presence of common mechanisms for irradiance encoding across more than 70 million years of independent evolution indicates strong evolutionary pressure for tight control of circadian rhythms by accurate reporting of environmental illumination to the brain. ■

REFERENCES AND NOTES

1. A. Liu *et al.*, *Science* **379**, 376 (2023).
2. K. Lucas, *The Conduction of the Nervous Impulse* (Longmans, 1917).
3. E. D. Adrian, *Basis of Sensation* (Hafner, 1928).
4. A. Huxley, *Biogr. Mem. Fellows R. Soc.* **43**, 13 (1997).
5. F. Rieke, D. Warland, R. de Ruyter van Steveninck, W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, 1997).
6. D. M. Berson, F. A. Dunn, M. Takao, *Science* **295**, 1070 (2002).
7. S. Hattar, H. W. Liao, M. Takao, D. M. Berson, K. W. Yau, *Science* **295**, 1065 (2002).
8. I. Provencio *et al.*, *J. Neurosci.* **20**, 600 (2000).
9. M. von Schantz, I. Provencio, R. G. Foster, *Invest. Ophthalmol. Vis. Sci.* **41**, 1605 (2000).
10. M. S. Freedman *et al.*, *Science* **284**, 502 (1999).
11. E. S. Milner, M. T. H. Do, *Cell* **171**, 865 (2017).
12. K. A. Roecklein *et al.*, *Neurosci. Biobehav. Rev.* **37**, 229 (2013).
13. K. B. Koronowski, P. Sassone-Corsi, *Science* **371**, eabd0951 (2021).

10.1126/science.adf9350

GRAPHIC: A. FISHER/*SCIENCE*

RETROSPECTIVE

Dieter Oesterhelt (1940–2022)

Pioneer of membrane protein research

By **Peter Hegemann**[1] *and* **Hartmut Michel**[2]

Dieter Oesterhelt, the biochemist who identified the first known microbial rhodopsin, died on 28 November 2022 at age 82. Dieter's discovery of the color and function of the so-called purple membrane of the halophilic bacterium *Halobacterium salinarum*—and the availability, simplicity, and stability of its single protein, bacteriorhodopsin—laid the foundations for the membrane protein research field that he launched in 1971. Microbial rhodopsins act as light-driven pumps, channels, and enzymes, making the field of optogenetics, in which light controls the activity of targets such as neurons, possible.

Born in 1940 in Munich, Dieter spent much of his career close to home. After receiving his diploma in chemistry in 1965 from the University of Munich, he remained there to work with biochemist and Nobel laureate Feodor Lynen on the structure of fatty acid synthetase, earning his PhD in biochemistry in 1967. Disappointingly, the crystals he obtained were disordered, so he joined the laboratory of Walther Stoeckenius, a biophysicist at the University of California, San Francisco, in 1969 to use electron microscopy. In 1979, after brief stints at the Friedrich Miescher Laboratory in Tübingen and the University of Würzburg, Dieter was appointed as a director at the Max Planck Institute of Biochemistry in Martinsried, near Munich, where he worked until his retirement in 2008.

While in San Francisco, Dieter met biophysicist Allen Blaurock, who was working on purple-colored membrane patches from *H. salinarum*. Dieter started the biochemical characterization of this strange membrane as a side project. He discovered that it contains only one protein and showed that its color is caused by the cofactor retinal, just as in animal rhodopsins. For that reason, he named the chromoprotein "bacteriorhodopsin" (BR).

Back in Munich, Dieter found that illumination bleached the purple color reversibly, similar to the color change of rhodopsin in animal eyes upon illumination. But the color of BR regenerated rapidly and did not require enzymatic reactions. In 1972, Dieter proposed that BR functions as a light-driven proton pump, constituting the second most abundant type of photosynthesis in the oceans. BR energizes halobacteria by creating an electrochemical proton gradient across the cytoplasmic membrane, which is used to drive the synthesis of adenosine triphosphate. This relation was the essence of Peter Mitchell's chemiosmotic theory, which BR experiments finally confirmed.

As a Max Planck director, Dieter carefully characterized the bioenergetic functionality of BR, studied its reconstitution from apoproteins, and analyzed the dynamics of the light-triggered proton-pumping reaction cycle. At a time when no membrane protein sequence, gene sequence, or structural information about membrane proteins was available, Dieter and his collaborators used x-ray crystallography, ultrafast spectroscopy, Fourier transform infrared spectroscopy, Raman spectroscopy, and nuclear magnetic resonance to conduct pioneering work on the structure and function of BR.



In 1978, Dieter's new field leapt forward when Har Gobind Khorana and Yuri Ovchinnikov each independently published the amino acid sequence of BR—the first known sequence of a membrane protein. Khorana sequenced the BR gene 2 years later. These two findings triggered a long process of molecular characterization of BR, led by a fruitful and friendly competition between Dieter and Khorana, who shared a deep mutual respect. Both labs, with their many collaborators, elucidated the molecular properties and constraints that determine BR's color, photodynamics, and constituent amino acids. The technology and molecular approaches that were developed for and with BR laid the groundwork for Dieter's and others' rapid characterization of myriad related microbial rhodopsins.

Dieter contributed to other fields as well. He discovered a plant ferredoxin in halobacteria, representing one of the earliest known cases of horizontal gene transfer. He and colleagues showed that the accessory chlorophyll of the photosynthetic reaction center is the first transient electron acceptor of the electron released by the chlorophyll special pair after photoexcitation. With great enthusiasm, Dieter supported diverse technical BR applications, including in biocomputing, holography, and pattern recognition.

Dieter's passion for science came through in his support for fellow scientists. He encouraged many former students and co-workers to launch independent research in his department; we were both grateful beneficiaries of this generosity. The success of Dieter's department arose from its creative atmosphere, where science came first but every individual was appreciated, respected, and supported. Dieter enjoyed lab parties and welcomed everyone to join him on excursions. An avid runner, tennis player, downhill skier, and mountaineer, he enjoyed being challenged athletically by his group members.

In addition to being an enthusiastic experimentalist, Dieter was an active science administrator and science policy adviser, always seeking to identify the best science that could be supported by the available funds. With geneticist Ernst-Ludwig Winnacker, Dieter established the Gene Center Munich, which served as a model for collaboration between the University of Munich and the Max Planck Society. Above all, he generously supported young colleagues as they made their transitions to independence.

For his groundbreaking science and scientific leadership, Dieter received many honors and awards. Already in 1982, he received the Feldberg Prize for Anglo-German Scientific Exchange. His pioneering achievements in biochemistry and molecular biology earned him the Otto Warburg Medal of the German Society for Biochemistry and Molecular Biology in 1991. In 2021, he shared the Albert Lasker Basic Medical Research Award for the development of optogenetics.

Dieter's collaborators and trainees around the world have lost a deeply respected and trusted adviser, supervisor, and friend, who appreciated both strengths and weaknesses but always pushed his colleagues forward into unexplored territories. The global scientific community has lost an exceptional scientist and visionary leader. We owe him a great deal and will continue to honor his memory in our work. ■

10.1126/science.adg5458

[1]Institute of Biology, Humboldt-Universität zu Berlin, Berlin, Germany. [2]Max Planck Institute of Biophysics, Frankfurt, Germany. Email: hegemann@hu-berlin.de

PHOTO: MAX PLANCK SOCIETY



Podcasts can spread scientific knowledge to diverse audiences but can also exacerbate echo chambers.

ESSAY

Science in the age of podcasts

A powerful 21st-century medium affords new opportunities for learning and creative thinking

By **Alex Gomez-Marin**

In his visionary 1964 book *Understanding Media: The Extensions of Man*, the Canadian media theorist Marshall McLuhan claimed that "the medium is the message," meaning that what is said is less consequential than how or where it is said (*1*). Accordingly, he argued, the study of human communication should give primacy to the communication medium itself.

If McLuhan is right, then podcasts are not only an atypical species in the media ecosystem but a potentially transformative one. The combination of message, messenger, and moment makes them an idiosyncratic medium for conceiving and communicating burgeoning ideas. This is therefore a timely moment to reflect on what this media form can offer to science and science communication writ large.

In 2019, in one of the first and most thorough studies of the science podcasting landscape, Lewis MacKenzie reported that the total number of such programs increased linearly from 2004 to 2010, followed by an exponential growth between 2010 and 2018 (*2*). And this was before the COVID-19 pandemic. Shows continue to increase in number, reaching growing audiences in many different countries, languages, and cultures.

The many merits of podcasts are obvious but nontrivial. Such programming is often cheap or even free, both to produce and to consume. And their digital nature allows podcasts to reach not only very large audiences but also diverse ones. Crucially, podcasting bolsters complexity and delivers pluralism. The intimate is made public in a bottom-up ecology of individual creators and self-organized communities.

In addition, and especially relevant to the dissemination of scientific knowledge, podcasts disentangle human communication from time and space. Conversations that might once have been held at meetings or specialist conferences can now be accessed from anywhere, helping to reduce the carbon footprint of knowledge diffusion, while sedimenting an invaluable online library for new learning. Hearing different voices discussing science in different ways may also help dispel outdated stereotypes of whom science is made by and for, while the informal, long-form, conversational nature of this type of medium encourages candid discussion that helps listeners understand the nuance of scientific knowledge making, bringing to the fore the perspectival and participatory nature of rational inquiry.

Of course, a dark side of podcasting exists too. We are all prey to clickbait content and audience capture dynamics, which can reinforce echo chambers. Content production for its own sake can be a pitfall too. Misinformation and disinformation, however, are arguably its most dangerous perils.

Should we attempt to filter out contentious "scientific" material? This leads to broader philosophical and sociopolitical questions. Censoring untrustworthy podcasts is likely to be not only fruitless but feckless. Follow "the science," some experts say, but science is not monolithic. Beyond compulsive accusations of conspiracy or collusion, podcasting can nurture an ethos of consilience. Curation, rather than censorship, will become a true art form in times to come.

So, what is really at stake? Why should scientists care? Not long ago, the paradigmatic way to produce and disseminate scientific knowledge was by letters and lectures. We now post preprints, attend virtual conferences, and download papers electronically. Drawing on notions advanced by American cultural anthropologist Margaret Mead regarding the generational making and remaking of society (*3*), we must realize, celebrate, and capitalize on the fact that early career and future scientists fall within a demographic that is the primary consumer of podcasts today (*4*). Podcasts are thus poised to play a central role in how these individuals make and take scientific knowledge. But podcasting's greatest virtue is probably its fitness to present the scientific enterprise as it actually is: an ongoing complex process of intersubjective consensus, rather than a machine for the making of unquestionable truths. ■

The author is at the Instituto de Neurociencias, Consejo Superior de Investigaciones Científicas–Universidad Miguel Hernández de Elche, Alicante, Spain.
Email: agomezmarin@gmail.com

REFERENCES AND NOTES

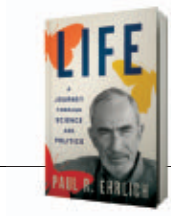
1. M. McLuhan, *Understanding Media: The Extensions of Man* (McGraw-Hill, 1964).
2. L. E. MacKenzie, *R. Soc. Open Sci.* **6**, 180932 (2019).
3. M. Mead, *Culture and Commitment: A Study of the Generation Gap* (Natural History Press, 1970).
4. "Podcast Industry Report: Market Growth and Advertising Statistics in 2022," Insider Intelligence, 29 July 2021; https://www.insiderintelligence.com/insights/the-podcast-industry-report-statistics/.

10.1126/science.adg5427

PHOTO: MIXETTO/ISTOCK.COM

A population pessimist turns 90

Paul Ehrlich reflects on life, love, and the science and politics of the human predicament



Life: A Journey Through Science and Politics
Paul R. Ehrlich
Yale University Press, 2023.
408 pp.

By **Peter Gleick**

For decades, the pioneering biologist and environmental scientist Paul Ehrlich, now 90, has published prolifically on topics ranging from species dynamics to population and consumption to nuclear war. In addition, he has provided straightforward explanations of science that have reached millions of people, including more than 20 appearances on *The Tonight Show* with Johnny Carson, who found Ehrlich to be brilliant and funny, with a sharp wit and acerbic tongue. Those attributes are all on display in Ehrlich's compelling autobiography, *Life: A Journey Through Science and Politics*.

Ehrlich's life has spanned the Great Depression, World War II, the Cold War, decades of stunning scientific advances, the flowering of civil rights, and a quadrupling of Earth's population. His memoir includes remarkable stories of his research, travels, friends, colleagues, and scientific controversies that still roil today. It is also a love story, highlighting Ehrlich's passion and esteem for both science itself and for his wife of nearly 70 years, Anne Ehrlich, an intellectual and artistic powerhouse in her own right and Ehrlich's long-time coauthor.

Ehrlich's earliest work focused on insect dynamics through field surveys that took him from the Arctic to tropical rainforests. He discovered new species of butterflies, studied tropical fishes and birds, and redefined the very concept of species, helping to create the field of population biology by integrating zoology, botany, and entomology with population dynamics, genetics, behavior, and evolution.

But the arc of Ehrlich's life has also included watching his beloved butterflies disappear as human populations have exploded, natural lands have been paved over, and climate change, pesticide use, and habitat destruction have expanded. This produced in him "a deep sadness—and a drive to reverse the ongoing destruction of the living world before it's too late" and fueled his efforts to raise the alarm about the risk of "a collapse of civilization as the existential threats—those endangering our very existence" worsen.

Ehrlich's reminiscences about students and colleagues are remarkably funny, insightful, loyal, and generous. But he does not suffer fools gladly, and his book includes pointed barbs at science deniers and "destructive segments of the American population" who reinforce the pathology of unlimited growth, kowtow to the fossil fuel industry and other corporate interests ("psychopathic engines of social and environmental destruction"), and fail to act on the world's problems.



Paul Ehrlich and an assistant catch checkerspot butterflies in California's Jasper Ridge Biological Preserve in 1960.

In 1968, the Ehrlichs wrote *The Population Bomb*, which shined a spotlight on the famine, social collapse, and environmental destruction that they believed would come to pass if population growth and other environmental threats continued to expand exponentially. It garnered massive public attention but was also attacked by free-market economists who assumed infinite growth is always good, optimists who believed technology can overcome adverse consequences of expanding populations, and some who criticized the book for failing to adequately acknowledge the role of consumption, wealth inequality, and racism in the population debate.

In *Life*, Ehrlich decries how his concerns about population growth were co-opted by racists who argued that the cure was to stop certain populations (almost always nonwhite) from reproducing and by those who turned a blind eye to consumption by the rich. Indeed, throughout the book, he reflects on the roles racism and sexism have played in his own experiences, including his efforts to organize sit-ins to desegregate restaurants as a graduate student in Kansas and to push back against the racist claims of eugenicist William Shockley, with whom he overlapped at Stanford University.

Modern understanding of the issues of population, inequality, consumption, and technology is far more nuanced than when *The Population Bomb* was first published. Ehrlich himself, together with Anne and other collaborators, helped define the intertwined role these factors play, but as he puts it, bringing down population numbers remains necessary to realizing "justice and equity and creating a world with any chance of sustainability."

Ehrlich still fears that we are headed for political and social unrest or even environmental collapse, but despite his deep pessimism, he acknowledges that we are making progress. Birth rates have dropped because of improved education, availability of contraceptives, information about family planning, and decreasing poverty. Action on climate change is accelerating. Concern for the environment is growing. If these trends continue, he writes, "the prospects for humanity would brighten considerably."

Life is also Ehrlich's reflection on the inevitability of his own death. He laments declining health and physical abilities and watching longtime friends and colleagues pass away, although "the thought of living forever has no appeal and being dead itself holds no fears." His greatest regret, he says, is not knowing the ultimate fate of humanity. Whatever our eventual fate, we are likely to be better off because of Ehrlich's remarkable contributions. ■

The reviewer is at the Pacific Institute, Oakland, CA 94612, USA. Email: pgleick@pacinst.org

10.1126/science.adg5246

PHOTO: PROPERTY OF PAUL EHRLICH



LETTERS

Edited by **Jennifer Sills**

EU deforestation law overlooks emerging crops

The new European Union (EU) deforestation law, for which a provisional political agreement has been reached (*1*), aims to limit imported commodities identified as the main drivers of deforestation, such as cocoa, coffee, oil palm, and soya (*2*). According to the current proposal, none of these products "would be allowed to enter or exit the EU market if they were produced on land subject to deforestation or forest degradation" after 31 December 2020 (*3*). Because production and consumption patterns can shift rapidly as global markets fluctuate, legislation should also allow for limits on emerging commodities that degrade forests; cashew-driven deforestation in the Afrotropics is a poignant example.

Cultivation of cashew (*Anacardium occidentale L.*) has grown seven-fold in recent decades, from 1.1 million ha in 1988 to 7.1 million ha in 2020, and demand continues to rise (*4*). As of 2020, the total area covered by cashew was still much lower than oil palm (29 million ha) and soybeans (127 million ha), but it has quickly reached the scale of coffee (11 million ha) and cocoa (12 million ha) (*5*). Furthermore, compared with cashew, coffee and cocoa are increasing at a much slower rate; since 1988, coffee has seen little change, and cocoa has doubled (*5*). Cashew, like other crops that have devastated tropical forests worldwide (*2*), is grown in full sun monocultures at vast scales. Although cashew research remains scant, there is mounting evidence that plantations offer little habitat for native species and encroach upon biodiverse forests inhabited by African forest elephants (*Loxodonta cyclotis*), chimpanzees (*Pan troglodytes*), and other endangered species (*6–10*).

The new deforestation law is an important step to prevent the EU-driven loss of forest, carbon stocks, and biodiversity, particularly in megadiverse tropical regions (*11*). However, to meet these goals, the list of commodities regulated by the EU law needs to build in the option to add emerging crops with high deforestation potential. Without that flexibility, crops such as cashew will remain unregulated as they increasingly drive forest destruction across the tropics.

Luke L. Powell[1,2,3,4]*, Joana Capela[1,2,5], Patrícia Guedes[1,2], Pedro Beja[1,2,5]

[1]Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO)–Rede de Investigação em Biodiversidade e Biologia Evolutiva (InBIO), Laboratório Associado, Universidade do Porto, 4485-6661 Vairão, Portugal. [2]BIOPOLIS Program in Genomics, Biodiversity, and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Portugal. [3]Biodiversity Initiative, Houghton, MI 49913, USA. [4]Institute of Animal Health and Comparative Medicine, University of Glasgow, Glasgow G128QQ, UK. [5]CIBIO–InBIO, Laboratório Associado, Instituto Superior de Agronomia, Universidade de Lisboa, 1349-017 Lisboa, Portugal.
*Corresponding author.
Email: Luke.L.Powell@gmail.com

REFERENCES AND NOTES

1. European Commission, "Green Deal: EU agrees law to fight global deforestation and forest degradation driven by EU production and consumption" (2022); https://ec.europa.eu/commission/presscorner/detail/en/ip_22_7444.

PHOTO: DE ROUVROY/NITIDAE, 2021



The effects of emerging crops such as cashew (shown here in Mozambique) could soon rival the deforestation caused by coffee and cocoa.

2. F. Pendrill *et al*., *Science* **377**, eabm9267 (2022).
3. European Commission, "Proposal for a Regulation of the European Parliament and of the council on the making available on the Union market as well as export from the Union of certain commodities and products associated with deforestation and forest degradation and repealing Regulation (EU) No 995/2010 COM/2021/706 final (Publications Office of the European Union)" (2021); https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0706.
4. United Nations Conference on Trade and Development, "Commodities at a glance: Special issue on cashew nuts" (2021); https://unctad.org/webflyer/commodities-glance-special-issue-cashew-nuts/.
5. Food and Agriculture Organization of the United Nations Statistics Division, "Crops and livestock products" (2022); https://www.fao.org/faostat/en/#data/QCL/visualize/. To see the area covered by the crops cited, go to the Download Data tab. In the first box, click the Regions tab. Select "World + (Total)." In the Elements box, select "Area harvested." In the Items box, select "Crops Primary > (List)." Click the arrow to the left to display all crops, and select "Cashew nuts, in shell," "Cocoa beans," "Coffee, green," "Oil palm fruit," and "Soya beans." In the Years box, select 1988 and 2020. At the bottom of the page, click the Show Data button.
6. F. Monteiro *et al*., *Sustainability* **9**, 1666 (2017).
7. S. Vasconcelos *et al*., *Biol. Conserv*. **191**, 779 (2015).
8. E. Bersacola, C. M. Hill, K. J. Hockings, *Sci. Rep.* **11**, 4569 (2021).
9. S. C. Pereira, C. Lopes, J. P. Pedroso, *Rem. Sens. Appl*. **26**, 100746 (2022).
10. D. Brugiere, I. Badjinca, C. Silva, A. Serra, M. Barry, *Oryx* **40**, 442 (2006).
11. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), "Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services," S. Díaz *et al*., Eds. (IPBES Secretariat, Bonn, Germany, 2019).

10.1126/science.adf9994

Hidden costs of Europe's deforestation policy

The European Union (EU) soon plans to implement a law that would limit the importation of products known to drive deforestation (*1*), including Brazilian products originating from deforested areas in the Amazon. The law's goal of removing incentives that lead to forest degradation is laudable, given the rate of deforestation (*2*). However, by focusing on forests alone, the law is potentially putting other ecosystems at greater risk, as well as overlooking the reliance of forests on adjacent regions. To prevent unintended consequences, the EU should expand the law to include ecosystems other than forests.

In response to the law, agribusinesses in the Amazon will likely search for a way to continue exporting their products to the EU. One strategy could be to move their operations to nearby regions that are not primarily forests, such as the Cerrado. Located adjacent to the Amazon, the Cerrado spans more than 2 million km² and is the most diverse savanna in the world (*6*). However, more than 46% of the land in the region has been converted to pasture (*6, 7*), and local water sources have been depleted (*8*). If agribusinesses moved to the Cerrado, they would convert additional land to agriculture or pastures, increasing vegetation loss (*3*–*5*). The resulting water and energy shortages (*5*, *6*, *8*) would affect everyone in the region, including agribusinesses (*9*), and potentially disrupt supply chains for Cerrado products such as almonds and pequi (*10*).

In addition to providing needed protection to ecosystems other than forests, broadening the EU law could better protect forests. The Amazon and the Cerrado interact through atmospheric water circulation (*10*, *11*). By indirectly causing further land degradation in the Cerrado, the law could disrupt the atmospheric dynamics that support biodiversity in the Amazon. To protect other ecosystems and the full spectrum of forests' needs, the EU should expand its policy as soon as possible, and legislation across the globe should aim to protect every type of vulnerable ecosystem.

Geraldo W. Fernandes[1,2]*, Hernani F. M. Oliveira[2,3], Helena G. Bergallo[2,4], Vitor N. T. Borges-Junior[2,5], Guarino Colli[2,3], Stephannie Fernandes[6], Nathan C. Fonsêca[2,7], Adrian A. Garda[2,8], Carlos E. V. Grelle[2,5], André V. Nunes[2,9], Lucas N. Perillo[1,2], Tainá C. Rocha[2,10], Domingos J. Rodrigues[2,11], Ricardo R. da Silveira-Filho[2,7], Helena Streit[2,12], Tiago S. P. Toma[1,2], Pedro L. Viana[2,10], Fabio O. Roque[2,9]

[1]Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil. [2]Brazilian Knowledge Center on Biodiversity, Belo Horizonte, MG, Brazil. [3]Universidade de Brasília, Brasília, DF, Brazil. [4]Universidade do Estado do Rio de Janeiro, Rio de Janeiro, RJ, Brazil. [5]Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil. [6]Florida International University, Miami, FL, USA. [7]Universidade Federal Rural de Pernambuco, Recife, PE, Brazil. [8]Universidade Federal do Rio Grande do Norte, Natal, RN, Brazil. [9]Universidade Federal de Mato Grosso do Sul, Campo Grande, MS, Brazil. [10]Museu Paraense Emílio Goeldi, Belém, PA, Brazil. [11]Universidade Federal de Mato Grosso, Cuiabá, MT, Brazil. [12]Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.
*Corresponding author.
Email: gw.fernandes@gmail.com

REFERENCES AND NOTES

1. European Commission, "Green Deal: EU agrees law to fight global deforestation and forest degradation driven by EU production and consumption" (2022); https://ec.europa.eu/commission/presscorner/detail/en/ip_22_7444.
2. "Deforestation in 2021 increased by 20%, with growth in all biomes," Mapbiomas (2021); https://mapbiomas.org/en/deforestation-in-2021-increased-by-20--with-growth-in-all-biomes-1.
3. E. K. H. zu Ermgassen *et al*., *Environ. Res. Lett*. **15**, 3 (2020).
4. F. Gollnow *et al*., *Environ. Res. Lett*. **17**, 11 (2022).
5. M. E. Ferreira *et al*., *Water* **14**, 22 (2022).
6. G. W. Fernandes *et al*., Cerrado: Em Busca de Soluções Sustentáveis (Vozes, Rio de Janeiro, (2016) [in Portuguese].
7. Y. Dou *et al*., *Geogr. Sci*. **28**, 11 (2018).
8. G. S. Hofmann *et al*., *Glob. Change Biol*. **27**, 17 (2021).
9. C. C. Pereira, G. W. Fernandes, *Science* **377**, 6603 (2022).
10. A. A. Rodrigues *et al*., *Glob. Change Biol*. **28**, 22 (2022).
11. N. Villoria *et al*., *Nat. Commun*. **13**, 5476 (2022).

10.1126/science.adg2983

Animal agriculture's potential financial risks

With a contribution of 16.5% to global greenhouse gas emissions (*1*), animal agriculture is a key driver of climate change, second only to fossil fuels. Soy production—three-quarters of which is used as livestock feed (*2*)—and beef production are the top two drivers of deforestation in the Amazon (*3*). Livestock farming poses a risk to more than 17,900 species listed on the International Union for Conservation of Nature's Red List of Threatened Species (*4*). Because the

environmental threats posed by animal agriculture can translate into a range of business risks for animal protein suppliers, environmentally responsible businesses are safer investments than those that cause harm (*5*). Therefore, businesses within the animal agriculture industry should measure and disclose their environmental impacts, and financial institutions should protect themselves from losses by investing in businesses that are good environmental stewards.

Business practices that degrade the environment can lead to business risks in the form of trade restrictions, stricter environmental regulations (for which compliance is costly), and litigation. The resulting financial losses can spill over to investors, who are then left with loan defaults and stranded assets (*5*). For example, the European Union (EU) recently agreed on a new law that bans the import of deforestation-linked commodities such as beef and soy. To maintain eligibility to import to the EU, suppliers face strict due diligence and traceability requirements with potential implications for investors (*6*). Similarly, New Zealand is set to become one of the first countries to tax greenhouse gas emissions from the animal agriculture sector (*7*). Meanwhile, American pork producer Smithfield Foods has been forced to pay millions in damages due to pollution from its animal waste lagoons (*8*).

Despite the potential financial risks resulting from environmental impacts, businesses and investors are not closely monitoring or attempting to mitigate the damage they cause. A recent survey of 60 of the largest meat, fish, and dairy producers with combined assets of US$324 billion found that 77% of them did not measure their greenhouse gas emissions or have targets in place to mitigate them. In addition, none of the 50 surveyed meat and milk producers had a comprehensive deforestation policy in place (*9*).

So far, the finance sector is not adequately taking environment-related financial risks into account when making investment decisions within the animal agriculture space. Between 2015 and 2020, 2500 investment firms, banks, and pension funds invested more than $478 billion in meat and dairy companies globally (*10*). Over the past decade, the International Finance Corporation (IFC), the private-sector lending arm of the World Bank, channeled more than $1.8 billion into mostly large-scale livestock and factory farming corporations (*11*).

Given that the World Bank Group is seen as a benchmark in environmentally responsible investment (*12*), it should address animal agriculture–related risks through its safeguard systems and exclusion lists. Currently, however, the World Bank and IFC only require their borrowers to improve resource and emissions efficiency (*13*), which may not be enough to prevent investments in environmentally—and thus financially—risky projects. Instead, the World Bank Group should establish strict investment criteria for large-scale livestock and factory farming operations. It should lead by example and avoid financing businesses that pose unacceptable greenhouse gas emission and deforestation risk.

Divya Narain
School of Earth and Environmental Sciences and Centre for Biodiversity and Conservation Science, University of Queensland, St Lucia, QLD, Australia. Email: divyanarain01@gmail.com

REFERENCES AND NOTES

1. R. Twine, *Sustainability* **13**, 6276 (2021).
2. W. Fraanje, T. Garnett, "Soy: Food, feed, and land use change," Foodsource: Building Blocks (2020).
3. World Wildlife Fund, "What are the biggest drivers of tropical deforestation?" (2018).
4. IUCN Red List, "The IUCN red list of threatened species" (2023); http://iucnredlist.org/search. On the left, click "Threats" to show all. Then click the arrow next to "2. Agriculture & Aquaculture." Under "2.3. Livestock farming and ranching" is the number of species affected: 17922.
5. FAIRR, "Factory farming: Assessing investment risks" (2016).
6. European Commission, "Green Deal: EU agrees law to fight global deforestation and forest degradation driven by EU production and consumption" (2022); https://ec.europa.eu/commission/presscorner/detail/en/ip_22_7444.
7. I. Smith, "Farmers protest 'unworkable regulations' of New Zealand's proposed farm levy," Euronews Green (2022); https://www.euronews.com/green/2022/10/20/cow-burps-to-be-taxed-under-world-first-proposals-by-new-zealand.
8. G. D. Robertson, "Court upholds hog verdict; Smithfield announces settlement," AP News (2020); https://apnews.com/article/north-carolina-courts-4b2f1db-4c21e03653851e81b81996410.
9. FAIRR, "Coller FAIRR protein producer index" (2019).
10. Feedback, "Butchering the planet: The big-name financiers bankrolling livestock corporations and climate change" (2020).
11. A. Mukpo, "World Bank's IFC pumped $1.8b into factory farming operations since 2010," Mongabay (2020).
12. S. Park, *Glob. Govern.* **13**, 535 (2007).
13. International Finance Corporation, "Performance standard 3: Resource efficiency and pollution prevention" (2012).

10.1126/science.abe2211

TECHNICAL COMMENT ABSTRACTS

Comment on "Interspecific competition limits bird species' ranges in tropical mountains"

Dingliang Xing, Jian Zhang, Fangliang He

Freeman *et al*. (Reports, 22 July 2022, p. 416) argue that interspecific competition rather than climate is the leading driver of bird species' elevational ranges. A reanalysis of their data shows no support for the competition hypothesis, but a strong effect of climate seasonality on species ranges. Their results are artifacts arising from a suboptimal model that misses important variables.

Full text: dx.doi.org/10.1126/science.ade2109

Response to comment on "Interspecific competition limits bird species' ranges on tropical mountains"

Benjamin G. Freeman, Matthew Strimas-Mackey, Eliot T. Miller

Xing *et al*. create new variables and fit models to argue against the hypothesis that interspecific competition shapes species' elevational ranges. However, their key newly created variable is best interpreted as a proxy for the important variable of the interspecific competition hypothesis. Thus, their reanalysis uncovers the patterns we already described that are consistent with the interspecific competition hypothesis.

Full text: dx.doi.org/10.1126/science.ade8043

NEXTGEN VOICES: SUBMIT NOW

The future of scientific societies

Add your voice to *Science*! Our new NextGen Voices survey for young scientists is open:

AAAS (the publisher of *Science*) turns 175 years old this year! AAAS's mission is to advance science, engineering, and innovation throughout the world for the benefit of all. Have scientific societies played a role in your career? Are scientific societies prepared to support scientists in a changing world? What is the most important role they could play, and what changes must they make to be more effective?

To submit, go to www.science.org/nextgen-voices

Deadline for submissions is 10 February. A selection of the best responses will be published in the 7 April issue of *Science*. Submissions should be 200 words or less. Anonymous submissions will not be considered. If your response is selected, *Science* can withhold your name from the published article upon request.

AAAS Annual Election: Preliminary announcement

Members will vote for president-elect and members of the Board of Directors

The AAAS election of general officers is scheduled to begin next month. All members will receive a ballot for election of the president-elect and members of the Board of Directors, who will begin their terms at the 2023 Annual Meeting. Biographical information for the candidates will be provided along with ballots. The general election slate is listed below.

As was approved by the Council on 22 July 2022, the newly constituted AAAS Section Committees will hold their full elections in the fall of 2023. In an effort to conserve resources, AAAS will be sending electronic election ballots only. In order to ensure that you receive your ballot, please make sure your email is up to date on your membership record. If you have any questions, please contact us at elections@aaas.org.

General Election Slate The following candidates will appear on the ballot for the position of **President-elect**:

Erich Jarvis, Professor, The Rockefeller University, and Investigator, Howard Hughes Medical Institute

Willie May, Vice President of Research and Economic Development, Morgan State University

The following candidates will appear on the ballot for the position of **Director**:

Rena D'Souza, Director, National Institute of Dental and Craniofacial Research, US National Institutes of Health

Kathleen Hall Jamieson, Elizabeth Ware Packard Professor of Communication and Director, Annenberg Public Policy Center, University of Pennsylvania

Jay Labov, Senior Advisor for Education and Communication (Retired), National Academies of Sciences, Engineering, and Medicine

Vassiliki Betty Smocovitis, Professor of History of Science, University of Florida

Constitution Amendment: Under the recommendation of the Governance Modernization Working Group and approval by the Council at the 22 July 2022 meeting, members will be asked to vote on an amendment to the Constitution as the final step toward approving the governance modernization efforts. For more information, visit aaas.org/amendment.



AAAS is dedicated to advocating for science and scientific evidence to be fully and positively integrated into public policy, and for the community to speak with one voice to advance science and engineering in the United States and around the world.

You can help us.

By making AAAS a beneficiary of your will, trust, retirement plan, or life insurance policy, you will become a member of our 1848 Society and will help fuel our work on behalf of science – including publishing the world's most promising, innovative research in the *Science* family of journals and engaging in the issues that matter locally, nationally and around the world.

If you plan to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

Legal Name: American Association for the Advancement of Science
Federal Tax ID Number: 53-0196568
Address: 1200 New York Avenue, NW, Washington, DC 20005

For more information or to let us know you are including AAAS in your future giving plans, contact the AAAS Office of Philanthropy and Strategic Partnerships at **1-800-215-1969** or **philanthropy@aaas.org.**

Additional details online at **www.aaas.org/1848Society.**

QUALITY CONTENT FOR THE GLOBAL SCIENTIFIC COMMUNITY

Multiple ways to stay informed on issues related to your research



Posters





Podcasts





Sponsored Collection Booklets





Advertorials





Webinars



Science
AAAS

Brought to you by the *Science*/AAAS Custom Publishing Office.



Scan the code and start exploring the latest advances in science and technology innovation!

Science.org/custom-publishing

RESEARCH

IN *SCIENCE* JOURNALS

Edited by
Michael Funk



COSMOCHEMISTRY

Nucleosynthetic anomalies in meteorites

The origins of the material that accreted to form Earth can be constrained using meteorites that contain leftover material from planet formation. Nucleosynthetic anomalies are small differences in isotope ratios left by incomplete mixing of presolar material. They are already known for refractory elements, which condense into dust grains first, but it has been unclear whether more volatile elements were fully mixed before planet formation. In two complementary papers, Martins *et al.* and Nie *et al.* identified nucleosynthetic anomalies in the moderately volatile elements zinc and potassium, respectively. They used cosmochemical models to determine that about 90% of Earth's mass was contributed by noncarbonaceous inner Solar System material, whereas about 10% came from carbonaceous outer Solar System material. The latter source contributed about 20% of Earth's potassium and half of its zinc. —KTS

Science, abn1021, abn1783, this issue pp. 369, 372

Isotopic analyses of meteorites, such as the one pictured here as a thin section, inform cosmochemical models of Earth's formation.

PHOTO: NICOLE XIKE NIE

CANCER

Limiting lymphopenia after glioblastoma

Lymphopenia, a lack of white blood cells, frequently occurs in patients with glioblastoma (GBM) after chemoradiation therapy, and its association with worse survival is not well understood. Analyzing peripheral blood from 20 patients treated for GBM, Ghosh *et al*. identified an elevated concentration of myeloid-derived suppressor cells (MDSCs) in patients with lymphopenia after treatment. The authors developed murine models to demonstrate a causal relationship between radiation-induced MDSCs and lymphopenia. Pharmacologic inhibition of MDSCs abrogated lymphopenia and improved survival in mice, suggesting that these classes of drugs may be a potential treatment to reduce lymphopenia in patients with GBM. —DLH

Sci. Transl. Med. **15**, eabn6758 (2023).

BIOSYNTHESIS

Limonoids in the limelight

Triterpenes are built from a common 30-carbon precursor, but through the action of tailoring enzymes can adopt a range of shapes and biological functions. De La Peña *et al*. investigated the biosynthesis of limonoids, which are found, for example, in citrus and neem plants, and characterized candidate genes involved in the production of two groups of these molecules. The authors identified 22 enzymes in total and reconstituted part of the pathways in a heterologous plant system. These results may facilitate the production of natural and unnatural limonoids with useful properties, such as plant-derived insecticides. —MAF

Science, adf1017, this issue p. 361

IMMUNOLOGY

Neutrophils cast NETs without dying

Neutrophil extracellular traps (NETs) bind to and kill microbes. Stojkov *et al*. found that, contrary to findings from previous studies, agonists that induced NET formation independently of multiprotein complexes called inflammasomes did not require the pore-forming protein gasdermin D (GSDMD) and did not cause cell lysis leading to cell death. Furthermore, canonical inflammasome activation caused GSDMD processing but was not required for NET release. These data suggest that the mechanisms leading to NET formation require further study and should be distinguished from those that lead to cell death. —JFF

Sci. Signal. **16**, eabm0517 (2023).

NANOPHOTONICS

Coupling at a distance

The merger of nanophotonics with quantum optics has provided a platform for the development of deterministic single-photon sources and

sources of entangled photons. Tiranov *et al*. show that multiple quantum emitters (two or three quantum dots) can be coupled through a photonic crystal waveguide. The waveguide helps to overcome the typically short-range nature of the dipole-dipole interactions and allows the signature properties of such a coupled system, subradiant and superradiant emission, to be observed. Control of such processes could provide an enabling step for scaling up deterministic solid-state photon-emitter interfaces and multiphoton-entangled sources for applications in quantum information processing. —ISO

Science, ade9324, this issue p. 389

EVOLUTIONARY ECOLOGY

Ecosystem-scale effects of rapid evolution

Plant traits influence many ecosystem processes, but the role of trait evolution is often overlooked in predicting ecosystem responses to environmental change. Vahsen *et al*. showed that rapid evolution in root traits of a dominant marsh sedge affected predictions of coastal wetland carbon dynamics. Growing seeds from different generations in a single location, they found considerable heritable variation in the allocation of belowground biomass. Input into an ecosystem model, this variation scaled up to substantial predicted differences in sediment accretion and marsh elevation, factors that influence a marsh's resilience to rises in sea level. —BEL

Science, abq0595, this issue p. 393



Large natural variation in the sedge *Schoenoplectus americanus* may aid in adaptation to environmental change.

SOLAR CELLS

Running hot and cold

Like other solar cells, commercial perovskite solar cells (PSCs) would not only need to maintain operation at the high temperatures generated in direct sunlight but also endure the lattice strain created by temperature changes throughout the year. Li *et al*. fabricated high-quality perovskite crystalline films by adding a fluorinated polymer, the dipoles of which lowered formation energy of the perovskite black phase, decreased defect density, and also tuned the surface work function for charge extraction. Power conversion efficiencies of 23% were achieved for 1-square-centimeter devices that retained over 90% of their efficiency after testing conditions for 3000 hours and after repeated cycling between −60° and 80°C. —PDS

Science, add7331, this issue p. 399

CRYSTALLINE ORDERING

Assembling aperiodic crystals

A metal-organic framework (MOF) has been made that contains an ordered lattice of connecting nodes but a network of aperiodically arranged linkers. Meekel *et al*. extended to MOFs the idea of Truchet tiles, in which a plane is covered with square tiles decorated with nonrotationally symmetric patterns to form aperiodic patterns. X-ray crystallography revealed that the zinc nodes formed an order lattice (see the Perspective by Yaghi and Rong). A comparison of x-ray diffuse scattering data with a three-dimensional difference pair-distribution function showed good agreement with a Truchet tile model. The aperiodicity created a disordered pore network within the crystal. —PDS

Science, ade5239, this issue p. 357; see also adf0261, p. 330

IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**

EVOLUTION

Of moth or bee

The white moon petunia, *Petunia axillaris*, has white flowers that absorb ultraviolet light and emit volatiles to attract the nocturnal hawkmoths that pollinate them. Conversely, the related species *Petunia secreta* displays magenta-colored, ultraviolet light–reflecting flowers that are visited by bees during the daytime. Luthi *et al.* show that deleting a single gene that regulates the biosynthesis of UV-absorbing flavonol pigments in *P. axillaris* turns the flowers pink and UV-reflective and makes the flowers attractive to both social and solidary bees. Therefore, single genes can cause major phenotypic effects that contribute to adaptation and speciation of the organisms. —DJ *Curr. Biol.* **32**, 5295 (2022).

A single gene determines whether petunia flowers are white and pollinated by night-flying hawkmoths, as shown here, or are pink and visited by day-flying bees.

SCIENTIFIC WORKFORCE

The ecology of publishing trends

Ecology, like most STEM disciplines, is overrepresented by male-identifying scientists from North America and Europe. As diversity efforts continue, Zettlemoyer *et al.* analyzed top-cited Ecology papers from 1960 to 2019 to look for evidence of increases in women authors from institutions outside of North America and Europe. Although the results showed more papers led by women authors over time, this corresponded with a decline in women as last authors, suggesting that women are not being retained in leadership positions. Despite ecology being a global science, authors from underrepresented countries were not found among top-cited publications. Interestingly, collaborative studies were found to increase over time, which can provide momentum for the inclusion of underrepresented collaborators as a way to form truly global, diverse collaborations. —MMc

Bull. Ecol. Soc. Am. 10.1002/bes2.2025 (2023).

SOLID-STATE MEMORIES

Coordinate framework memories

Materials that can be programmed with electrical pulses to change their conductance state can be used for memory applications. Robinson *et al*. show that the lithium-doped ruthenium analog of Prussian blue, ruthenium hexacyanoruthenate, can undergo four orders of magnitude changes in conductivity with voltage pulses. Compared with other Prussian blue compounds, the ruthenium analog had a greater degree of charge delocalization and conductivity between the Ru(II) and Ru(III) sites along

PHOTOS (TOP TO BOTTOM) MATT LAVIN/FLICKR; EPSILONERIDANI/WIKIMEDIA COMMONS



pathways that connect across the pores and less-conductive paths along the shorter nearest-neighbor directions. This material was inkjet printed onto flexible substrate to create electrochemical random access memory devices that were compatible with neurons and could switch conductance states through redox reactions with the neurotransmitter dopamine. —PDS

Adv. Mater. 10.1002/adma.202207595 (2022).

GENE THERAPY

Gene therapy can be skin deep

Dystrophic epidermolysis bullosa is a debilitating skin condition associated with poor wound healing caused by a deficiency in the collagen gene *COL7A1*. The result is blistering, scarring, and a variety of serious complications. Potential interventions for this disease are being investigated, but none has been approved to date. Guide *et al*. conducted a double-blind, intrapatient-randomized, placebo-controlled trial of a topical gene therapy in 31 patients with dystrophic epidermolysis bullosa. The *COL7A1* gene was delivered in a replication-defective herpesvirus vector, which is large enough to accommodate the target gene and specifically delivers it to the skin. For each patient, the medical team selected two similar wounds for direct comparison and demonstrated much greater chances of healing after gene therapy than after placebo treatment. —YN

N. Engl. J. Med. **387**, 2211 (2022).

CELL BIOLOGY

Caveolin-1 dolines for mechanoadaptation

Cells are constantly exposed to different types and intensities of mechanical forces. To accommodate these changes, they modulate their physical properties and adapt their plasma membrane (PM). Caveolae are small membrane invaginations that contribute to mechanoadaptation by flattening out, thereby buffering PM tension changes. Lolo *et al*. used experimental and computational approaches to study the role of caveolin-1 (Cav1), a component of caveolae, in mechanoprotection. Cav1 confers deformability and mechanoprotection independently from caveolae through modulation of PM curvature. Cav1-dependent PM remodeling stabilized noncaveolar invaginations called dolines that respond to low and medium mechanical forces. Dolines conferred mechanoprotection to cells devoid of caveolae. Restricting the size of Cav1 dolines hindered membrane buffering except at relatively high forces able to flatten caveolae. Thus, caveolae and dolines help cells adapt to a variety of mechanical stimuli. —SMH

Nat. Cell Biol. 10.1038/s41556-022-01034-3 (2022).

MATERIALS SCIENCE

A continuous color gradient

Structural coloration occurs when the structural features of an object are the same size as the wavelengths of visible light. When the structures are crystalline, the observed color will change with viewing angle and orientation. Schöttle *et al.* made a photonic colloidal glass using a controlled emulsion extraction process that synthesizes latex particles with a gradient in size depending on the reaction time. The particles were removed using a thin tube to retain laminar flow and assembled into films using a filtration technique. Because each particle size will scatter light of a different wavelength, the gradient films show vivid structural colors along their thickness and superior light scattering across the gradient. —MSL

Adv. Mat. 10.1002/adma.202208745 (2022).

STROKE

Rejuvenating immune cells

Aging is associated with worse outcome after stroke. It is also known that aging microglia, the primary innate immune cells in the brain, impair recovery from stroke in mice. Arabaizar-Rovirosa *et al.* investigated this connection and found that intracellular lipids accumulated in the microglia of aged mice and displayed dysfunctional immune responses after stroke. If microglia were depleted and this was followed by repopulation with new cells, then the inflammation and neurological damage associated with stroke were ameliorated in older animals. It appears that aged microglia might contribute to worse stroke outcomes in the elderly. —MM

EMBO Mol. Med. e17175 (2022).

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

IMMUNOLOGY

Rethinking innate immunity activation

Multicellular organisms initiate innate immune responses to microbial infections through the activation of pattern recognition receptors (PRRs). These receptors recognize ligands generated by microbes, and they are largely considered to respond to a successful and productive infection. However, PRR ligands are often microbial proteins and nucleic acids that would not normally be exposed during infection. In a Perspective, Kagan proposes that PRRs in plants and animals actually recognize infection infidelities, and mistakes during infection elicit an innate immune response. Rethinking how PRRs are activated could open a new avenue for anti-infective drug development. —GKA

Science, ade9733, this issue p. 333

TROPICAL FOREST

Losing the Amazon

The Amazon rainforest is a biodiversity hotspot under threat from ongoing land conversion and climate change. Two Analytical Reviews in this issue synthesize data on forest loss and degradation in the Amazon basin, providing a clearer picture of its current status and future prospects. Albert *et al*. reviewed the drivers of change in the Amazon and show that anthropogenic changes are occurring much faster than naturally occurring environmental changes of the past. Although deforestation has been widely documented in the Amazon, degradation is also having major impacts on biodiversity and carbon storage. Lapola *et al*. synthesized the drivers and outcomes of Amazon forest degradation from timber extraction and habitat fragmentation, fires, and drought. —BEL

Science, abo5003, abp8622, this issue p. 348, p. 349

VACCINES

Location, location, location

To generate protective immune responses, vaccine antigens are designed to mimic native protein structures on foreign pathogens. After vaccination, antigens are transported from the injection site by lymphatic vessels to the draining lymph nodes, tissues where immune responses are initiated. Aung *et al*. found that in order to mount a successful antibody response, incoming antigens need to be localized to specific sites within the lymph nodes called follicles (see the Perspective by Kalia and Hickman). Those antigens that were not rapidly directed to follicles were destroyed by proteases. Vaccines that resulted in speedy delivery of antigens to follicles preserved intact antigen and generated better antibody responses. These findings may have implications for the design of vaccines against difficult to neutralize or genetically variable pathogens. —PNK and SMH

Science, abn8934, this issue p. 350; see also adf5134, p. 332

RESPIRATORY COMPLEXES

Complex I inhibition by biguanides

Metformin is a biguanide drug widely used for the treatment of type 2 diabetes, but its effects on metabolism are complex, and its simple structure and relatively weak binding have led to a variety of proposed mechanisms of action. Bridges *et al*. determined the structures of bovine complex I bound to a biguanide inhibitor that is related to metformin but binds more tightly. The structure revealed potential binding sites at three different locations, all of which could contribute distinct effects. The strongest site appeared to be the coenzyme Q–binding channel, and the authors propose that the primary mode of inhibition may be halting reactivation of the resting deactive state of the enzyme. —MAF

Science, ade3332, this issue p. 351

NEUROSCIENCE

Encoding illumination in primates

Light is an important environmental factor for most organisms because it regulates many physiological functions and behaviors. Central to this regulation is a photoreceptive molecule called melanopsin and its actions within the intrinsically photosensitive retinal ganglion cells (ipRGCs), neurons that send signals from the retina to the brain. Little is known about the ipRGCs of primates. Liu *et al*. developed a method of identifying primate ipRGCs and used it to investigate how they sense the visual world (see the Perspective by Martin). They found that ipRGCs encode the overall intensity of environmental illumination using specializations that are distributed across scales of biological organization. These cells thus show insensitivity to image detail and a wide dynamic range. —PRS

Science, ade2024, this issue p.376; see also adf9350, p. 335

GEOCHEMISTRY

A global rate for weathering

Chemical weathering is an important way that carbon dioxide is regulated over geological time scales. However, many factors go into the rate of weathering, making scaling from laboratory to global estimates challenging. Brantley *et al*. assessed weathering rates across several different length scales to understand the rate-limiting factors for major rock types on Earth (see the Perspective by Hilton). The authors used this information to upscale globally to get an estimate of the temperature dependence of chemical weathering. The resulting model will be helpful in understanding the weathering response to climate change in attempts to increase weathering in the future. —BG

Science, add2922, this issue p. 382; see also adf3379, p. 329

IMMUNODEFICIENCY

Rewired IRF4 activity

The transcription factor interferon regulatory factor 4 (IRF4) is critical for the development of the immune system, especially B and T cells. The IRF4 International Consortium identified seven patients from six families with a combined immunodeficiency who harbored a recurrent heterozygous mutation in IRF4 called p.T95R. Despite having normal levels of IRF4, patients with $IRF4^{T95R}$ and knock-in mice bearing the same point mutation had multiple B cell defects, including impaired differentiation and immunoglobulin secretion. B cells expressing $IRF4^{T95R}$ failed to express the genes required for plasma cell differentiation because of altered DNA-binding specificity involving a loss of binding to canonical interferon-stimulated response elements and acquired binding at noncanonical sites. This study defines a previously undescribed cause of immunodeficiency driven by an IRF4 mutation leading to distinct changes in its activity. —CO

Sci. Immunol. **8**, eade7953 (2023).

Science
JOURNALS AAAS



Publish your research in the *Science* family of journals

The *Science* family of journals (*Science, Science Advances, Science Immunology, Science Robotics, Science Signaling*, and *Science Translational Medicine)* are among the most highly-regarded journals in the world for quality and selectivity. Our peer-reviewed journals are committed to publishing cutting-edge research, incisive scientific commentary, and insights on what's important to the scientific world at the highest standards.

Submit your research today!

Learn more at **Science.org/journals**

REVIEW SUMMARY

TROPICAL FOREST

Human impacts outpace natural processes in the Amazon

James S. Albert*, Ana C. Carnaval, Suzette G. A. Flantua, Lúcia G. Lohmann, Camila C. Ribas, Douglas Riff, Juan D. Carrillo, Ying Fan, Jorge J. P. Figueiredo, Juan M. Guayasamin, Carina Hoorn, Gustavo H. de Melo, Nathália Nascimento, Carlos A. Quesada, Carmen Ulloa Ulloa, Pedro Val, Julia Arieira, Andrea C. Encalada, Carlos A. Nobre

BACKGROUND: The Amazon is a critical component of the Earth climate system whose fate is embedded within that of the larger planetary emergency. The Amazon is the most species-rich subcontinental-scale ecosystem and is home to more than 10% of all named plant and vertebrate species, concentrated into just 0.5% of Earth's surface area. The Amazon rainforest is also a critical component of the Earth climate system, contributing about 16% of all terrestrial photosynthetic productivity and strongly regulating global carbon and water cycles.

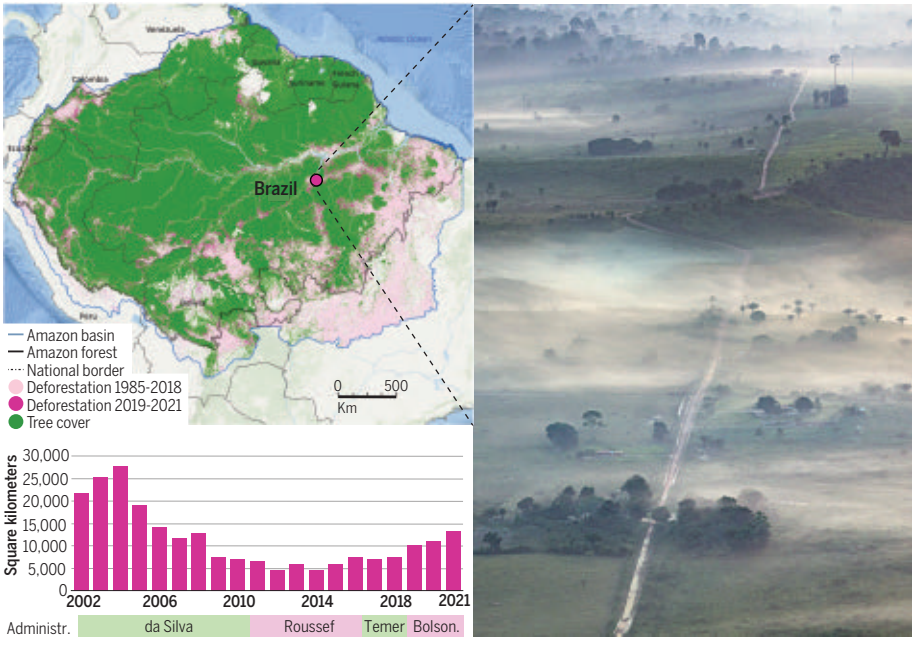
Amazonian ecosystems are being rapidly degraded by human industrial activities. A cumulative total of 17% of the original forest have already been cleared, and 14% replaced, by agricultural land use. After millions of years serving as an immense global carbon pool, under further warming the Amazon rainforest is predicted to become a net carbon source to the atmosphere. Some regions have already made the transition, with forest respiration and burning outpacing forest photosynthesis.

ADVANCES: In this Review, we compare rates of anthropogenic and natural environmental changes in the Amazon and South America and in the larger Earth system. We focus on deforestation and carbon cycles because of their critical roles on the Amazon and Earth systems. Data for South America were compiled for the Science Panel for the Amazon (SPA) Assessment Report, which details the many dimensions of the Amazon as a regional entity of the Earth system. The SPA report, coauthored by 240 scientists from 20 countries, documents epoch-scale transformations in Amazonian biodiversity, ecosystem function, and cultural diversity.

We found that rates of anthropogenic processes that affect Amazonian ecosystems are up to hundreds to thousands of times faster than other natural climatic and geological phenomena. These anthropogenic changes reach the scale of millions of square kilometers within just decades to centuries, as compared with millions to tens of millions of years for evolutionary, climatic, and geological processes. The main drivers of Amazonian habitat destruction and degradation are land-use changes (such as land clearing, wildfires, and soil erosion), water-use changes (such as damming and fragmenting rivers and increased sedimentation from deforestation), and aridification from global climate change. Additional important threats come from overhunting and overfishing, introduction of invasive exotic species, and pollution from the mining of minerals and hydrocarbons.

OUTLOOK: Given the outsized role of the Amazon in our planetary hydrological cycle, large-scale deforestation of this region is expected to push the whole Earth system across a critical threshold to a qualitatively different global climate regime. Quite aside from biodiversity losses, such a transformation will have multifarious and catastrophic consequences for human welfare, including widespread water and food insecurity that will lead to mass migrations and political instability. The key message is that Amazonian environments are being degraded by human industrial activities at a pace far above anything previously known, imperiling its vast biodiversity reserves and globally important ecosystem services.

The Amazon is now perched to transition rapidly from a largely forested to a nonforested landscape, and the changes are happening much too rapidly for Amazonian species, peoples, and ecosystems to respond adaptively. Policies to prevent the worst outcomes are known and must be enacted immediately. We now need political will and leadership to act on this information. To fail the Amazon is to fail the biosphere, and we fail to act at our peril. ■



Amazon deforestation is accelerating from a combination of anthropogenic drivers, including drier climatic conditions and policies that favor industrialized agriculture. (Top left) Map of Amazon showing location of wildfires, 1985 to 2021. (Right) Recently burned primary forest near Rurópolis, State of Pará, Brazil, 17 September 2020. (Bottom left) Rate of deforestation in the Brazilian Amazon is now rising rapidly under environmental policies of the Bolsonaro administration. After millions of years serving as an immense global carbon pool, the Amazon rainforest is becoming a net carbon source to the atmosphere.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: jalbert@louisiana.edu
Cite this article as J. S. Albert *et al.*, *Science* 379, eabo5003 (2023). DOI: 10.1126/science.abo5003

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abo5003

CREDITS: (MAP) ESRI, GARMIN-GEBCO, NOAA NGDC; (PHOTO) MARIZILDA CRUPPE/AMAZONIA REAL

REVIEW

TROPICAL FOREST

Human impacts outpace natural processes in the Amazon

James S. Albert[1]*, Ana C. Carnaval[2], Suzette G. A. Flantua[3], Lúcia G. Lohmann[4], Camila C. Ribas[5], Douglas Riff[6]†, Juan D. Carrillo[7], Ying Fan[8], Jorge J. P. Figueiredo[9], Juan M. Guayasamin[10], Carina Hoorn[11], Gustavo H. de Melo[12], Nathália Nascimento[13], Carlos A. Quesada[14], Carmen Ulloa Ulloa[15], Pedro Val[16,17,18], Julia Arieira[19], Andrea C. Encalada[20], Carlos A. Nobre[21]

Amazonian environments are being degraded by modern industrial and agricultural activities at a pace far above anything previously known, imperiling its vast biodiversity reserves and globally important ecosystem services. The most substantial threats come from regional deforestation, because of export market demands, and global climate change. The Amazon is currently perched to transition rapidly from a largely forested to a nonforested landscape. These changes are happening much too rapidly for Amazonian species, peoples, and ecosystems to respond adaptively. Policies to prevent the worst outcomes are known and must be enacted immediately. We now need political will and leadership to act on this information. To fail the Amazon is to fail the biosphere, and we fail to act at our peril.

The Amazon is a critical component of the Earth climate system, whose fate is embedded within that of the larger planetary emergency. Along with the two polar ice sheets and coral reefs, the Amazon [as defined in (*1*)] is one of four major ecosystems of the Earth system that are rapidly approaching or surpassing the threshold to a qualitatively degraded state (*2*, *3*). The Amazon is by far the most species-rich subcontinental-scale ecosystem, being home to more than 10% of all named plant and vertebrate species concentrated into just 0.5% of Earth's surface area (*4*). Yet Amazonian biodiversity is grossly underestimated, with perhaps only about 10% of the species yet described (*5*). Amazonian biodiversity is the evolutionary source for much of the world's plants and animals (*6*, *7*), serving as the core of a biogeographic realm that hosts about one-third of all known species on Earth (*8*).

The Amazon is also a crucial provider of global ecosystem services, contributing about 16% of all terrestrial photosynthetic productivity (*9*) and strongly regulating global carbon and water cycles (*10*, *11*). Yet global warming is rapidly increasing climate variability in the Amazon. Extreme droughts and record floods have occurred in nine of the past 15 years, compared with just four extreme droughts and three record floods in the previous century (*11*). These extreme weather events are substantially lowering the threshold for wildfires at the rainforest margins; altering biogeochemical cycles; and leading to widespread deforestation, habitat degradation, and wetland loss (*9*, *12*).

Given the outsized role of the Amazon in our planetary hydrological cycle, large-scale deforestation threatens to push the whole Earth system across a critical threshold to a qualitatively different global climate regime (*13*). Quite aside from biodiversity losses, such a transformation will have multifarious and catastrophic consequences for human welfare, including widespread water and food insecurity (*14*–*16*), leading to mass migrations and political instability (*16*).

In this Review, we compare rates of anthropogenic and natural environmental changes in the Amazon and other regions of South America and also compare these rates with other processes in the larger Earth system. Data for South America were compiled from the Science Panel for the Amazon (SPA) Assessment Report (*1*), which details the many dimensions of the Amazon as a regional entity of the Earth system. The SPA report—coauthored by 240 scientists from 20 countries, including members of Indigenous Peoples and Local Communities (IPLCs)—documents epoch-scale transformations in Amazonian biodiversity, ecosystem function, and cultural diversity. The report also summarizes the major social and ecological transformations of the Amazon through human history and presents sustainable development pathways for the Amazon into the near future. The key messages of this Review are that multiple strong changes to the Amazon being driven by modern human activities are happening far too fast for the survival of its species and ecosystems (*17*) and that widespread Amazon deforestation would be an irreversible catastrophe for the global climate system (*9*, *18*).

Amazon in motion

The Amazon is perched to transition rapidly from a largely natural to degraded and transformed landscape, under the combined pressures of regional deforestation and global climate change (*19*, *20*). As of 2019, a cumulative total of about 17% of the pre-Columbian Amazon forest had been cleared, and 14% replaced, by human agriculture landscapes—89% for pasture and 11% for crops (*21*). After millions of years serving as an immense global carbon pool, under further warming the Amazon rainforest is predicted to become a net carbon source to the atmosphere [for example, (*22*, *23*)]. Some parts of the Amazon have already made the transition, with forest respiration and burning outpacing forest photosynthesis (*24*).

As we enter the third decade of the 21st century, portions of the southern and eastern Amazon are changing to a disturbance-dominated regime (*25*, *26*). Under global drivers of climate change, much of the Amazon is experiencing pronounced increases in the frequency and severity of floods, droughts, and wildfires (*12*, *27*). The basin-wide impacts of landscape desiccation have far surpassed the variability of natural hydrological and biogeochemical cycles since the start of the current climate epoch, the Holocene, ~11,700 years ago (*28*). Further, several other ecologically and biodiversity-rich regions of the Neotropics outside of the Amazon (such as the Atlantic Rainforest or Mata Atlântica, Caatinga, Cerrado, Chocó, and Puna) are also facing accelerating threats from modern human activities (*1*, *7*).

Before the Anthropocene (starting around 1950), the Amazon had maintained natural humid and tropical environments, including forests and wetlands, over most of lowland northern South America for tens of millions of years (*4*). Amazonian ecosystems have persisted

[1]Department of Biology, University of Louisiana at Lafayette, Lafayette, LA, USA. [2]Department of Biology and Ph.D. Program in Biology, City University of New York (CUNY) and CUNY Graduate Center, New York, NY, USA. [3]Department of Biological Sciences, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway. [4]Universidade de São Paulo, Instituto de Biociências, Departamento de Botânica, São Paulo, SP, Brazil. [5]Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, AM, Brazil. [6]Instituto de Biologia, Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brazil. [7]Department of Biology, University of Fribourg and Swiss Institute of Bioinformatics, Fribourg, Switzerland. [8]Department of Earth and Planetary Sciences, Rutgers, The State University of New Jersey, NJ, USA. [9]Institute of Geoscience, Center of Mathematical and Earth Sciences, Universidade Federal Rio de Janeiro, RJ, Brazil. [10]Instituto Biósfera, Laboratorio de Biología Evolutiva, Universidad San Francisco de Quito USFQ, Quito, Ecuador. [11]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, Netherlands. [12]Department of Geology, Federal University of Ouro Preto, Ouro Preto, MG, Brazil. [13]Institute of Advanced Studies, University of São Paulo, SP, Brazil. [14]Coordination for Environmental Dynamics, National Institute for Research in Amazonia, Manaus, AM, Brazil. [15]Missouri Botanical Garden, St. Louis, MO, USA. [16]School of Earth and Environmental Sciences, Queens College, CUNY, New York, NY, USA. [17]Ph.D. Program in Earth and Environmental Sciences, CUNY Graduate Center, New York, NY, USA. [18]Department of Geology, Federal University of Ouro Preto, Ouro Preto, MG, Brazil. [19]Science Panel for the Amazon (SPA), São José dos Campos, SP, Brazil. [20]Instituto Biósfera, Universidad San Francisco de Quito, Quito, Ecuador. [21]Institute of Advanced Studies, University of São Paulo, SP, Brazil.
***Corresponding author. Email: jalbert@louisiana.edu**
†Present address: Ecoinformatics Studio, Rio de Janeiro, RJ, Brazil.

through many profound climatic and evolutionary transformations, including the formation and draining of inland seas and mega-wetlands during most of the Miocene (~23 million to 10 million years ago), and transitioned into a fluvial landscape in the late Miocene to Pliocene (~10 million to 2.3 million years ago) (*29*), alternated ice-age and interglacial climates during the Pleistocene (~2.6 million to 0.01 million years ago) (*29*, *30*), and shifted land-use practices of Indigenous peoples during the Holocene (*31*).

Thus, quite unlike the expansive temperate and boreal forests of the northern hemisphere, which were repeatedly cleared and pushed southward by low temperatures and continental glaciers during the Pleistocene and then regenerated in the Holocene, Amazonian rainforests have never previously confronted regional-scale deforestation (*32*, *33*). This ecosystem persistence over evolutionary time scales resulted in the Amazon becoming both a center and source of biodiversity for the whole Neotropical region (*6*, *34*).

In the Amazon, more than in most other regions, forest-rainfall feedback is required to maintain the current forest cover (*35*). About half of the precipitation over the Amazon is recycled from evapotranspiration, with about 14.1 trillion cubic meters of water per year falling as precipitation over the whole basin, compared with the Amazon River discharge of about 7.3 trillion cubic meters per year. Amazonian forest cover buffers the ecosystem against variations in precipitation and fire (*36*, *37*). This dependence of the state of the system on its history (hysteresis) is a common feature of many ecological systems at large spatial and temporal scales, in which the observed state of a system cannot be predicted on the basis of current conditions alone.

Amazon forest extent and structure is therefore highly sensitive to widespread forest degradation and removal (*38*, *39*). Clearcutting parts of the Amazon forest exposes the landscape to an irreversible regime shift, from a forested to a nonforested landscape, with a wide range of deleterious consequences (*12*, *40*). Beyond a certain threshold, deforestation and regional aridification will become locked in a vicious cycle that drives a runaway transformation of lush rainforests to degraded savanna-like agricultural landscapes (*25*, *41*).

Drivers of Amazon destruction and degradation

The main regional-scale drivers of Amazonian habitat destruction and degradation arise from land-use changes (such as deforestation, wildfires, or soil erosion), water-use changes (such as damming and fragmenting rivers, increased sedimentation from deforestation, pollution from the mining of minerals and hydrocarbons, or ground-water extraction), and aridification from global climate change (*5*, *18*). The main effects of climate change today are precipitation changes, and sea level rise will likely have major effects in the near future. Overhunting and overfishing (*42*), the introduction of invasive exotic species (*43*), and pollution (*44*) are additional important threats to biodiversity and ecosystem function at local to regional scales in the Amazon and other ecosystems. Here, we focus on deforestation and carbon cycles because of their critical roles on the Amazon and Earth systems.

The most rapid environmental changes in the Amazon today are driven by land converted from forests and degraded pastures into soy and livestock production, primarily for export (*45*, *46*). By 2019, about 867,000 km^2 or about 14% of the Amazon forest had been cleared, especially in the Brazilian states of Pará, Mato Grosso, Rondônia and Amazonas, in order of greatest contribution to deforestation (*21*). Between 1995 and 2017, 17% of the Amazon rainforest was degraded by logging, fire, windthrow, or road expansion (*47*). Under the auspice of globalization, Amazonia is being integrated into global commodities markets, mostly soybean, beef, and timber (*48*).

The immediate crisis is driven by the logging and burning of closed-canopy tropical rainforests to clear land for agriculture and pasture. Agricultural expansion is the leading cause of regional deforestation worldwide and in South America (*49*, *50*). The legal construction of roads, dams, and other infrastructure, combined with many illegal activities (such as forest clearcutting, logging and burning, mining, illicit crops, and clandestine roads) have driven the agricultural frontier deep into the Amazon margins over the past 20 years (*51*, *52*). During this same period, soybean exports from Brazil to China surged by 2000%, primarily as animal feed to supply rapidly increasing meat consumption in China, and South America is currently the largest source of biomass imports to the European Union (*53*).

The great soybean plough-up of South America during the early 21st century is the farthest outlier of anthropogenic changes from the regression lines for South America in Fig. 1. This landscape transformation is roughly comparable in total area and proportion of landscape surface with other regional-scale "great plough-ups" of history, such as the spread of grain culture across monsoon Asia from about 3000 to 1000 years ago; the Northern European plains from about 1500 to 1000 years ago; the Russian Steppes in the 18th and 19th centuries; the Great Plains of North America in the late 19th and early 20th century; and the ongoing expansion of palm oil plantations in Indonesia, Malaysia, and many other countries.

Effective forest-protection policies act by removing the international financing of market-driven land-conversion projects. Two of the largest funding sources are the Inter-American Development Bank (IDB), based in Washington, DC (*54*), and the Belt and Road Initiative (BRI) of the Chinese government. The Initiative for the Integration of the Regional Infrastructure of South America (IIRSA) is a massive infrastructure program of road and dam construction launched in 2000. Most IIRSA environmental impacts derive from road construction in the Brazilian states of Amazonas and Acre and the Colombian states of Caquetá and Guaviare, providing increased access for accelerated expansion of beef production, oil extraction, and mining (*55*).

BRI-financed hydroelectric and water-diversion projects are planned to dredge and canalize hundreds of river kilometers in Ecuador and Perú (*56*). BRI-supported water diversion projects will expand soybean cultivation on more than 74,000 km^2 and hydrologically link Amazonian tributaries to neighboring drainages. Once completed, these projects will convert major southern tributaries (such as Tapajos and Xingu rivers) into a network of artificial reservoirs with poorly known but negative impacts to local biodiversity and IPLC livelihoods and the function of regional hydrological systems (*57*).

The effectiveness of forest-protection policies has varied over the past 20 years (*52*, *58*). The Action Plan for the Prevention and Control of Deforestation in the Legal Amazon (PPCDAm), launched in 2004, improved the deforestation

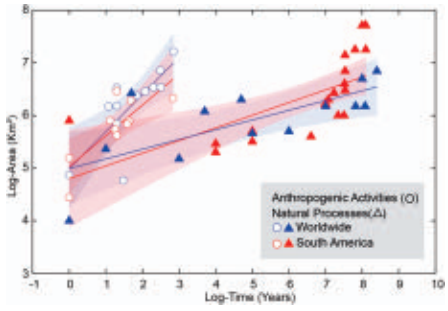


Fig. 1. Temporal and spatial scales of anthropogenic and natural processes in the Earth system. Data for 55 cases, with references in Table 1. Circles and triangles indicate anthropogenic and natural processes, respectively; red and blue symbols indicate processes from South America and globally, respectively. All regressions are power functions represented as linear curves on a log-log plot. Anthropogenic South America (n = 10 activities), y = 106,443 × 0.5853, coefficient of determination (R^2) = 0.2455. Anthropogenic global (n = 12 activities), y = 96,870 × 0.7071, R^2 = 0.8214. Natural South America (n = 21 activities), y = 102,364 × 0.185, R^2 = 0.4565. Natural global (n = 13 activities), y = 97,678 × 0.1849, R^2 = 0.4669. Anthropogenic processes occur at rates several orders of magnitude faster than those of natural processes.

monitoring system, reinforced environmental inspections, and promoted land tenure for IPLCs in legally protected areas. These actions were strengthened over time by the Soy Moratorium (from 2006) and the Black List of municipalities with highest deforestation rates (from 2008). Together, these actions substantially reduced access of industrial farming interests to international markets and financial credit (*53*, *58*). Whereas the Temer and Bolsonaro administrations (2016–2022) undermined the PPCDAm, weakened the new Brazilian Forest Code, dismantled environmental agencies, and suppressed the Sugarcane Agroecological Zoning Act of 2009 (*59*), markedly increasing deforestation rates, one of the first acts of the new Lula administration was to reestablish the PPCDAm.

Global climate change represents the other imminent threat to the Amazon and other ecosystems, affecting forest dynamics, carbon and nutrient cycling, and freshwater and coastal ecosystems (*60*, *61*). As predicted by climate models (*62*, *63*) and well documented by climatic records (*11*), precipitation patterns are becoming more variable in time and space, with more frequent and severe floods (*64*) and more persistent and widespread droughts (*39*). Climate change is rapidly desiccating the southern and eastern portions of the Amazon rainforest, contributing to higher frequency and severity of wildfires and contraction of the southern forest margin. Concomitant sea level rise is projected to inundate the biodiverse floodplain and coastal mangroves and estuaries, converting them to nearshore marine habitats and threatening coastal livelihoods (*65*).

How fast is the Amazon changing?

We compiled age and area estimates for 55 different anthropogenic and natural processes affecting terrestrial and aquatic ecosystems in South America and globally, including 11 anthropogenic and 21 natural processes in the former and 13 and 11 processes in the latter (Table 1). Ensemble rates were assessed by the exponent value of power-function regressions applied to each of these four categories.

We found that rates of anthropogenic processes affecting Amazonian ecosystems are up to hundreds to thousands of times faster than they are for natural climatic and geological phenomena (Fig. 1). These anthropogenic changes have reached the scale of millions of square kilometers within just decades to centuries, as compared with millions to tens of millions of years for evolutionary, climatic, and geological processes. Destruction of Amazonian environments is far outpacing species', ecological interactions', and ecosystems' capacity to respond adaptively (*32*, *66*). The rate at which modern human activities is driving extinctions in the Neotropics is between 1000 and 10,000 times higher than the natural or "background" rate as estimated from the fossil record (*17*, *67*).

These anthropogenic changes to Amazonian environments are coupled to processes worldwide, racing ahead many times faster than those of natural counterbalancing processes in the Earth system (*68*). Among the most important ongoing imbalances are accelerating rates of climate change (*69*), sea level rise (*70*), terrestrial vegetation turnover (*32*), river delta avulsion (*71*), tropical deforestation (*72*, *73*), extinction (*74*), and soil erosion and waterway sedimentation (*75*–*77*). Whereas the residence time of carbon through the atmosphere, hydrosphere, and lithosphere is on the order of millennia to millions of years, modern human extraction and burning of fossil fuels occurs at time frames of decades to centuries (*78*). Global climate changes during the most recent deglaciation (for example, the Pleistocene-Holocene transition) occurred on the time frame of centuries to millennia as compared with ongoing anthropogenic changes that are observed at a decadal scale (*79*).

Given the key role of the Amazon in the Earth system, the causes and consequences of Amazonian and global system degradation are strongly linked (*1*), and the pace of anthropogenic changes exceeds that of many natural processes at regional to global scales (Fig. 1). For example, average annual global deforestation over the past decade has exceeded afforestation by about 100,000 km^2, causing a net loss of forest of about 1.4% every year (*80*). Global soil erosion exceeded soil formation by 35.9 billion tons (Gt) in 2012, representing a 2.5% increase over the erosion estimate from 2001 (*81*). Rates of vegetation change equal or exceed the deglacial rates globally, indicating that the scale of human effects on terrestrial ecosystems now exceeds the massive vegetation transformations during the most recent major global climate change event (*32*). In the Amazon, changes in the precipitation patterns, because of deforestation or withdrawal, are having a strong impact on the frequency and magnitude of intermittency of rivers and streams specially in the southeastern part of the Amazon. Last, although accurate data on groundwater withdrawals are difficult to collect, estimates indicate that depletion far exceeds recharging in most parts of the world, with net losses of up to 20% per year in some highly populated and aridifying regions of North America and Asia (*82*).

Global consequences of Amazon degradation

From a climate perspective, widespread Amazon degradation would be an irreversible global catastrophe. Amazonian forests and soils contain about 180 ± 30 Gt of carbon (GtC); approximately half of this carbon is stocked in the form of vegetation biomass, and the other half remains as soil carbon stocks (*9*). By comparison, this Amazonian carbon volume is equivalent to about 26% of the 690 ± 80 GtC released into the atmosphere by all human activities since the Industrial Revolution (1750 to 2020), achieved primarily by burning fossil fuels and land-use changes (*83*). Anthropogenic carbon emissions during this time period raised atmospheric carbon dioxide (CO_2) from 277 to 415 parts per million (ppm) and increased the average global temperature to 1.2°C above preindustrial levels. Releasing all the Amazonian carbon into the atmosphere would initially increase the airborne CO_2 concentration by an additional 85 ppm, representing another concerning ~0.5°C increase (*83*).

Under the 2015 Paris Climate Accords, to keep atmosphere warming below 2°C global civilization cannot emit more than 465 Gt more carbon, and the Amazon alone contains about 32 to 44% of that carbon total. Yet Amazonian fires from 2010 to 2018 released about 0.5 to 1.5 GtC per year into the atmosphere, whereas forest growth during this time period removed only about 0.5 GtC per year (*84*). The approximately 4.5 to 9.0 GtC left in the atmosphere is similar to total carbon emissions of Japan during this interval, which ranked fifth among nations for carbon pollution (*85*). To better compare the volume of Amazon carbon impact on global climate, we note that Amazonian afforestation in the centuries after the Iberian conquest (around 1500 to 1700) captured about 7.4 GtC (3.5 ppm CO_2 equivalent) from the atmosphere, perhaps contributing to the global cooling episode known as the Little Ice Age (*86*).

The adverse consequences of global anthropogenic carbon emissions extend beyond the Amazon to the whole Earth system. Without sufficient abatement, melting polar ice sheets will contribute more than 13 m (~43 feet) to global sea level rise by 2500, with complete loss of the Earth's ice sheets projected within the next 400 to 700 years (*87*). Ongoing melting of the Western Antarctic is projected to fragment the Thwaites Eastern Ice Shelf within the next five years, raising sea levels by more than 0.6 m and destabilizing neighboring glaciers (*88*, *89*). In an ice-free world, global sea levels would reach ~65 m (~213 feet) above the present level, as high as they were in the supergreenhouse world of the Eocene about 56 million years ago (*90*). Such melting would raise the global sea level 93 to 162 mm per year averaged over the next few centuries, starting slow (averaging 3.1 mm per year in the past 30 years) and accelerating toward the final collapse of the ice sheets. By comparison, sea levels rose about 60 m during the early and mid-Holocene (11,700 to 7000 years ago), at an average rate of about 12.9 mm per year (*91*). Thus, the potential anthropogenic rate of sea level rise in the next few years and decades is more than seven times faster than the maximum

Table 1. Anthropogenic and natural processes affecting terrestrial and aquatic ecosystems. Data specific to the Amazon is indicated with an asterisk. LIP, large igneous provinces; E-O, Eocene-Oligocene.

Category	Process	Age (million years)	Area (km^2)	References	Notes
Anthropogenic global	Land equipped for irrigation: 1700–2020	320	3,442,500	(*136*, *137*, *158*)	
	Wetland loss: 1700–2009	309	7,220,000	(*159*)	
	Freshwater withdrawals: 1800–2000	200	3,443,500	(*160*, *161*)	
	Land equipped for irrigation since 1900	120	2,863,500	(*136*, *137*, *158*)	
	Land equipped for irrigation since 1950	70	2,383,500	(*136*, *137*, *158*)	
	Urban land expansion: 1970–2000	30	58,000	(*162*)	
	Land equipped for irrigation since 2000	20	703,500	(*136*, *137*, *158*)	
	Urban land expansion: 2010–2030	20	1,527,000	(*162*)	Most likely forecast
	Habitat loss from agricultural expansion: 2020–2050	20	3,350,000	(*69*)	
	Global forest cover loss: 2000–2012	12	1,500,000	(*163*)	Forests with >50% tree cover
	Global deforestation: 2012	1	74,532	(*163*)	Forests with >50% tree cover
Anthropogenic South America	Marine incursions to 80 M: by 2700	680	2,125,900	(*164*)	Area estimated from maps using ImageJ
	Rangeland desertified South America: 1960–2008	48	1,943,000	(*165*)	Area estimated from claim of 30% loss
	Amazon deforestation* 1975–2018	43	788,353	(*20*)	
	Petroleum concessions*: 1970–2008	38	688,000	(*166*)	Western Amazon (n = 188 concessions)
	Soybean expansion South America: 2000-2019	20	2,870,000	(*52*)	
	Soybean expansion Amazon*: 2000–2019	20	420,000	(*52*)	
	Anthropogenic forest loss: 2000–2017	18	540,000	(*26*)	
	Amazon fires*: 2003–2015	13	800,000	(*167*)	
	Amazon fires*: 2019	1	156,000	(*168*)	
	Amazon deforestation peak*: 2004	1	27,772	(*72*)	
Natural global	LIP: Siberian Traps	252,000,000	7,000,000	(*169*)	
	LIP: Ontong Java Plateau	120,000,000	1,500,000	(*151*)	
	Megariver captures stream orders 8 to 10	100,000,000	5,642,282	(*34*)	
	LIP: Deccan Traps	66,000,000	500,000	(*170*)	
	Megariver captures stream orders 6 to 8	10,000,000	253,195	(*171*)	
	Megariver captures stream orders 4 to 6	1,000,000	11,362	(*171*)	
	1 km bolide impacts	50,000	5,000	(*152*)	1 km diameter crater
	10 m bolide	500	2,150	(*152*)	Tunguska event, area deforested
	2.5 m bolide	50	1,875	(*152*)	Area deforested
Natural South America	Origins modern rainforest floras and faunas Western Gondwana	125,000,000	51,447,500	(*4*)	Western Gondwana is South America, Africa, and Arabia
	Megathermal forests across South America	125,000,000	17,840,000	(*4*)	
	Final separation South America and Africa	100,000,000	51,447,500	(*4*)	
	Diversification of modern rainforest floras and faunas	64,000,000	17,840,000	(*4*)	
	E-O global cooling, contraction of rainforests to tropical latitudes	34,000,000	14,000,000	(*4*)	
	Separation Amazon and Atlantic biotas, seasonally dry diagonal	34,000,000	7,000,000	(*4*)	
	Marine regression, expansion lowland basins	34,000,000	3,000,000	(*4*)	
	GAARlandia	33,000,000	4,000,000	(*4*)	
	Megariver captures in sub-Andean foreland	32,000,000	1,000,000	(*4*)	
	Pebas megawetland system	22,000,000	1,000,000	(*4*)	
	Expansion of C4 grasses and mammalian grazers	17,000,000	2,690,000	(*4*)	South American savannas
	Separation cis- and trans-Andean lowland biotas	12,000,000	2,000,000	(*4*)	Trans-Andean lowlands
	Desertification at continental periphery	10,000,000	1,708,000	(*4*)	Patagonia, Atacama, Sechura, Goajira, and Caatinga
	Great Amazonian Biotic Interchange (GAzBI)*	10,000,000	1,600,000	(*152*)	
	Rise of Fitzcarrald arch*	4,000,000	400,000	(*172*)	
	Ice ages cycles: forest-savanna*	100,000	500,000	(*173*)	
	Irion cycles: várzeas*	100,000	460,000	(*174*)	
	Irion cycles: igapos*	100,000	320,000	(*174*)	
	Megafauna extinctions—changes woody-savanna cover	10,000	290,000	(*174*)	
	Ice ages cycles: shorelines	10,000	200,000	(*164*)	

recorded rate after the most recent global deglaciation.

The rapid pace of human activities is readily seen in Stommel diagrams that plot the characteristic temporal and spatial scales of disparate human economic, geological, climatological, and biological processes (Fig. 2). In this context, it is useful to compare the modern anthropogenic biodiversity and climate crises with the Paleocene-Eocene Thermal Maximum (PETM) event, a global but relatively brief hyperthermal episode that occurred about 55.5 million to 54.5 million years ago. During the PETM, atmospheric CO_2 rose to the highest levels of the Cenozoic Era, and the global average temperature spiked about 5° to 8°C to a temperature about 9° to 14°C warmer than today, driving large changes to the geographic ranges and adaptive traits of many terrestrial and marine organisms (*92*). By contrast, current rates of change in CO_2 and global average temperature are hundreds of times faster than were during the PETM (*93*, *94*). Such unprecedentedly high rates of environmental change constitute the most important challenges to adaptation and persistence of plant and animal species in Amazonian ecosystems and to global civilization (*95*).

Transformative pathways for sustainable development

The current state and future fate of the Amazon are inextricably bound to that of the entire Neotropical region, the global biosphere as a whole, and the future of civilization worldwide (*45*, *48*, *96*). Preserving Amazonian biodiversity and ecosystem services will require fundamental changes to legal, economic, and energy systems at both regional and global scales. Policy actions must be implemented to reverse climate change and reduce economic incentives in the international trade system that support export-driven economic development (*97*). These changes to international legal and economic systems must deliberately be built into the next phase of the Anthropocene, when civilization transitions from carbon-based to renewable energy technologies and a bioeconomy of healthy standing forest and flowing rivers with sustainable governance (*98*, *99*).

A new legal framework

Successful economic development in many parts of the world has historically rested on a robust legal framework that incentivizes prosocial—and disincentivizes antisocial—behaviors and activities (*100*–*102*). Recent advances in environmental ethics and international justice provide robust legal standing for natural entities such as landscape features (rivers and forests) and nonhuman species (*103*, *104*). For example, in a landmark ruling, the Constitutional Court of Ecuador applied the constitutional provision on the "Rights of Nature" to safeguard cloud forests from mining concessions (*4*, *105*). This legal precedent was grounded in decades of scholarship (*106*, *107*), and similar laws have been codified in other countries (*98*, *108*). "Earth system law" provides a complementary approach for addressing gaps in governance that arise from improper deregulation and dispersed regulatory architecture across institutions and geographic regions (*25*, *109*). These legal tools can be designed to impose criminal penalties of heavy fines and imprisonment to criminalize activities that wantonly and substantially damage or destroy Amazonian ecosystems or that harm the health and well-being of Amazonian species (*110*, *111*). The importance of legal mechanisms in landscape preservation is well demonstrated by the success of the PPCDAm in reducing deforestation in Brazil from 2004 to 2015 and by decisions made at the federal level not to prosecute illegal activities that dramatically accelerated deforestation from 2016 to 2022 (*112*).

A new Amazonian bioeconomy

The sustainable use of biodiversity resources is an important path for developing Amazonian economies to become integrated into the international economy under advantageous conditions (*99*). More than 40 million people inhabit the Amazon region, with more than 65% living in urban areas, all of whom are affected by climate change. IPLCs play a critical role in shaping, protecting, and restoring ecosystems, biodiversity, and cultural diversity in the Amazon (*113*, *114*). A successful bioeconomy extends beyond extractive and export-based economic activities (such as lumber, mining, soy, and cattle) by prioritizing and monetizing biodiversity and ecosystem services and promoting broad development goals in education, health, sanitation, and employment. Improving the quality of life of the Amazonian population—in urban, peri-urban, and rural areas—is one of the principles of a bioeconomy based on standing forests and flowing rivers.

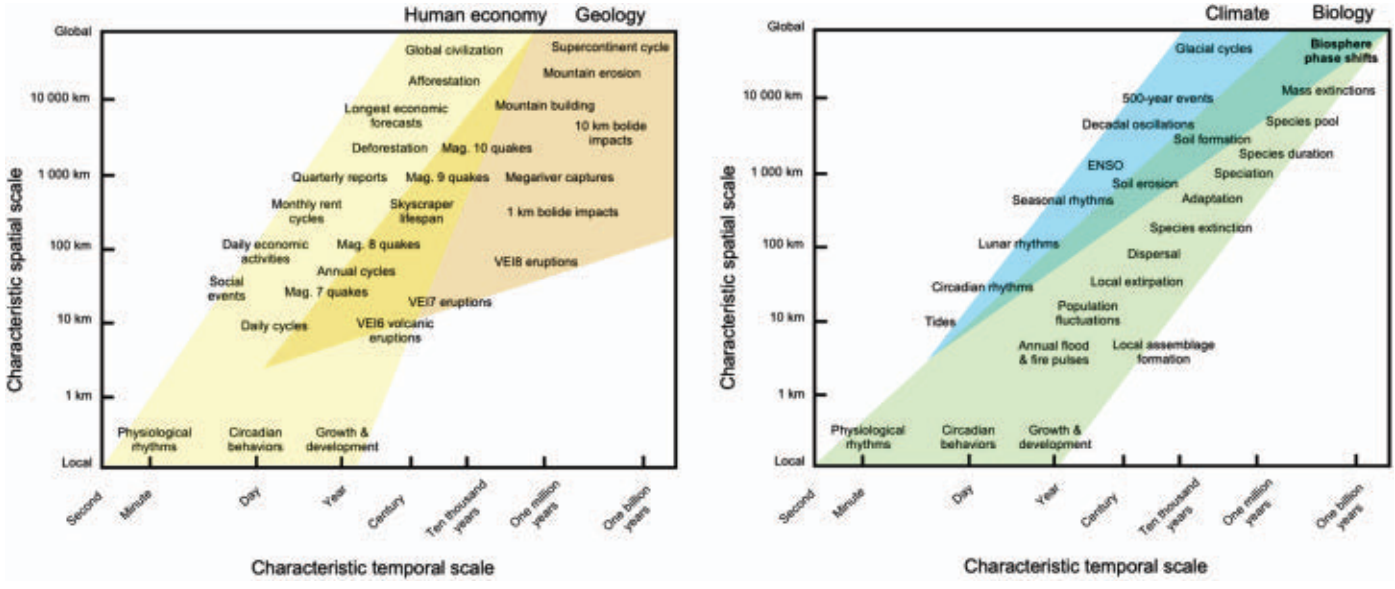


Fig. 2. Stommel diagrams estimating the temporal and spatial scales for 52 natural processes across four domains. (**Left**) Human economy (*73*, *76*, *77*, *138*–*142*) and geology (*143*–*152*). (**Right**) Climate (*81*, *153*, *154*) and biology (*155*–*157*). Axes are plotted by using logarithmic scales, with log seconds on the horizontal axis and log kilometers on the vertical axis. Biosphere phase shifts (top right) include long-wave climate (greenhouse-icehouse) cycles and distinct events such as the Neoproterozoic formation of an oxidizing atmosphere, Cambrian explosion of animal body plans, Devonian colonization of the continents and formation of terrestrial biotas, and the Anthropocene climate and biodiversity crises. Human economic activities affect larger spatial scales more rapidly than do most other natural processes.

Desired outcomes of a new Amazonian bioeconomy optimize carbon sequestration, biodiversity recovery, and human livelihoods (*115*, *116*). Sustainable bioeconomic development projects are most effective when they integrate modern scientific and commercial resources of urban communities with the traditional knowledge and skills accumulated by Indigenous and local farming communities over many generations (*48*). Lasting sustainability means prolonged coexistence of natural and human economic and social systems, and Amazonian development projects must therefore meet the immediate and long-term needs of the Amazonian population. Paramount among these needs are high-quality communication and transportation services to improve the commercialization of products, as well as institutional investments and international collaborations that support education, science, and technology institutions located within the Amazon. The installation of any new large-scale infrastructure projects (such as mega-dams or transportation arteries exceeding 500 km) must be avoided and replaced with low-impact alternatives (*117*). Mining initiatives that threaten Indigenous lands, the health of all Amazonian inhabitants, and biodiversity should also be avoided.

Resilient planning and management of Amazonian bioresources must necessarily prioritize the social and political actions that preserve species, habitat diversity, and functional redundancy; manage connectivity and feedback that stabilize longer-term processes over decades; promote reciprocal cultural and educational exchanges; and enhance integrated and decentralized (versus hierarchical and centralized) governance (*117*–*119*). Rates of deforestation in the Amazon since 2000 have closely responded to policy changes enacted at the national level that affect these kinds of social and political actions (*117*, *119*).

By contrast, market mechanisms based on international commodity pricing have entirely failed to assess the real economic and social values of Amazonian landscape and ecosystem resources (*99*, *120*). Further, prospects are dim for using market forces in landscape conservation efforts in the near future (*51*). Public policies to correct these market failures are available, modeled from strategies successfully used in other regions of the world where standing forests and flowing rivers have been allowed to persist for multiple decades, even under the context of intensive economic development (*121*, *122*). These policies successfully price the full market value of ecosystem services, provide incentives for activities that support forest and river preservation, and impose penalties for predatory and negligent actions (*123*).

The "Grand Energy Transition"

Preserving Amazonian biodiversity and ecosystem services requires modifying economic incentives in the international trade system that drive export-driven development (*97*). Such a Grand Energy Transition is already well underway (*124*); the average cost per unit energy for renewable energies has fallen below that of fossil fuels in aggregate for the first time in human history (*125*). Yet the barriers to complete this transition remain high, including the high costs of infrastructure installation and resistance by powerful stakeholders of the carbon economy (*126*). One of the biggest challenges is the high volume of fossil carbon still sequestered within the lithosphere; about 60% of oil and fossil methane gas and 90% of coal must be left in the ground to limit global warming to 1.5°C (*127*).

Yet time is running short. Emerging technologies, social innovations, and broader shifts in cultural practices are being implemented to support a resilient biosphere and help maintain a healthy Amazon (*95*, *128*). These shifts can be accelerated with economic and legal actions that support a post-carbon global economy that includes alternative energies, CO_2 capture and sequestration, and possibly geoengineering. New socioeconomic innovations must prioritize circular economic supply and waste networks and nurture green values and land ethics. New political and ecological innovations require coordination among leaders from the local, regional, and national levels. Widespread public support for greener development has already had qualitative impacts in many settings, and public awareness must be increased in Amazonian countries to influence elections and political decisions concerning environmental protection (*129*).

Policy actions and priorities

Long-term (decades to centuries) conservation critically relies on economic and legal support to Amazonian universities, research institutions, and scientific collections. These academic institutions are singularly situated to document Amazonian systems at multiple structural, geographic, and temporal scales and to characterize poorly known organisms (such as plants, fungi, invertebrates, and microbes), which are the "ecosystem engineers" that regulate biogeochemical cycles in Amazonian soils and surface and ground waters. These institutions also provide the skilled labor force required to monitor Amazonian environments through time and to train the next generation of Amazonian scientists.

Yet action is also required at broader scales. The global community must work closely and swiftly with national governments whose sovereignty includes Amazonian territory to enact economic, legal, and scientific actions that limit global warming to 1.5°C above preindustrial levels (*130*) and disincentivize activities for commodity export, especially soy, beef, timber, mineral, and hydrocarbon extraction (*131*). These actions are abstracted from the SPA Assessment Report (*1*, *132*) and other recent global environmental assessments (*133*, *134*). These actions recognize the knowledge and rights of IPLCs, who play a critical role in shaping, protecting, and restoring ecosystems and biodiversity in the Amazon and other tropical regions (*25*, *131*, *132*).

The most effective conservation actions enhance legal protections and punish illegal activities for areas under public, private, community, and Indigenous management, and reward companies, agencies, and communities committed to sustainable economic practices (*132*, *135*–*137*). These actions prioritize partnerships with IPLCs, areas with distinctive and threatened species, ecosystems, culturally important landforms, and the highest anthropogenic threat—those with the most rapidly expanding human footprint. International financial institutions (such as IDB and BRI) must immediately suspend funding for IIRSA mega-infrastructure projects (such as roads, bridges, railways, dams, ports, and mines) in Amazonia, pending thorough, independent, and regional-scale environmental assessments (*135*). Annual commodity supply chain reports of imports by country will enhance accountability. Success critically relies on robust, long-term partnerships among Amazonian people in the business, scientific, and IPLC communities. These partnerships provide sustained administrative, financial, and legal resources to IPLCs to secure land tenure rights; monitor, protect, and restore Amazonian ecosystems and biodiversity; and exchange biodiversity and conservation information between academic and local knowledge bases.

As we approach an irreversible tipping point for Amazonia, the global community must act now. Policies to prevent the worst outcomes have been successfully identified, but implementation is a matter of leadership and political will. To fail the Amazon is to fail the biosphere, and we fail to act at our own peril.

REFERENCES AND NOTES

1. Science Panel for the Amazon (SPA), *Amazon Assessment Report 2021* (United Nations Sustainable Development Solutions Network, 2021).
2. I. Amigo, When will the Amazon hit a tipping point? *Nature* **578**, 505–507 (2020). doi: 10.1038/d41586-020-00508-4; pmid: 32099130
3. M. H. Costa *et al.*, in *Amazon Assessment Report 2021*, C. Nobre *et al.*, Eds. (United Nations Sustainable Development Solutions Network, 2021).
4. J. M. Guayasamin *et al.*, in *Amazon Assessment Report 2021*, C. Nobre *et al.*, Eds. (United Nations Sustainable Development Solutions Network, 2021).
5. E. Berenguer *et al.*, in *Amazon Assessment Report 2021*, C. Nobre *et al.*, Eds. (United Nations Sustainable Development Solutions Network, 2021).
6. A. Antonelli *et al.*, Amazonia is the primary source of Neotropical biodiversity. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6034–6039 (2018). doi: 10.1073/pnas.1713819115; pmid: 29760058
7. A. S. Meseguer, P. O. Antoine, A. Fouquet, F. Delsuc, F. L. Condamine, The role of the Neotropics as a source of world tetrapod biodiversity. *Glob. Ecol. Biogeogr.* **29**, 1565–1578 (2020). doi: 10.1111/geb.13141

8. P. H. Raven *et al*., The distribution of biodiversity richness in the tropics. *Sci. Adv.* **6**, eabc6228 (2020). doi: 10.1126/sciadv.abc6228; pmid: 32917691
9. Y. Malhi *et al*., in *Amazon Assessment Report 2021*, C. Nobre *et al.*, Eds. (United Nations Sustainable Development Solutions Network, 2021).
10. M. Jung *et al*., Areas of global importance for conserving terrestrial biodiversity, carbon and water. *Nat. Ecol. Evol.* **5**, 1499–1509 (2021). doi: 10.1038/s41559-021-01528-7; pmid: 34429536
11. M. H. Costa, L. Borma, P. M. Brando, J. A. Marengo, S. R. Saleska, L. v. Gatti, in *Amazon Assessment Report 2021*, C. Nobre *et al.*, Eds. (United Nations Sustainable Development Solutions Network, 2021).
12. X. Xu, G. Jia, X. Zhang, W. J. Riley, Y. Xue, Climate regime shift and forest loss amplify fire in Amazonian forests. *Glob. Chang. Biol.* **26**, 5874–5885 (2020). doi: 10.1111/gcb.15279; pmid: 32662146
13. W. Steffen *et al*., Trajectories of the Earth System in the Anthropocene. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8252–8259 (2018). doi: 10.1073/pnas.1810141115; pmid: 30082409
14. R. B. Larson, *Just Add Water: Solving the World's Problems Using its Most Precious Resource* (Oxford Univ. Press, 2020); https://academic.oup.com/book/33629.
15. J. Marotzke, D. Semmann, M. Milinski, The economic interaction between climate change mitigation, climate migration and poverty. *Nat. Clim. Chang.* **10**, 518–525 (2020). doi: 10.1038/s41558-020-0783-3
16. D. J. Kaczan, J. Orgill-Meyer, The impact of climate change on migration: A synthesis of recent empirical insights. *Clim. Change* **158**, 281–300 (2019). doi: 10.1007/s10584-019-02560-0
17. V. Radchuk *et al*., Adaptive responses of animals to climate change are most likely insufficient. *Nat. Commun.* **10**, 3109 (2019). doi: 10.1038/s41467-019-10924-4; pmid: 31337752
18. C. A. Nunes *et al*., Linking land-use and land-cover transitions to their ecological impact in the Amazon. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2202310119 (2022). doi: 10.1073/pnas.2202310119; pmid: 35759674
19. C. A. Boulton, T. M. Lenton, N. Boers, Pronounced loss of Amazon rainforest resilience since the early 2000s. *Nat. Clim. Chang.* **12**, 271–278 (2022). doi: 10.1038/s41558-022-01287-8
20. D. C. da Cruz, J. M. R. Benayas, G. C. Ferreira, S. R. Santos, G. Schwartz, An overview of forest loss and restoration in the Brazilian Amazon. *New For.* **52**, 1–16 (2021). doi: 10.1007/s11056-020-09777-3
21. MapBiomas Amazonia, Collection 2.0 of annual maps of land cover, land use and land use changes between 1985 to 2018 in the Pan-Amazon (2020); https://amazonia.mapbiomas.org.
22. W. Hubau *et al*., Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature* **579**, 80–87 (2020). doi: 10.1038/s41586-020-2035-0; pmid: 32132693
23. M. J. P. Sullivan *et al*., Long-term thermal sensitivity of Earth's tropical forests. *Science* **368**, 869–874 (2020). doi: 10.1126/science.aaw7578; pmid: 32439789
24. L. v. Gatti *et al*., in *Amazon Assessment Report 2021*, C. Nobre *et al.*, Eds. (United Nations Sustainable Development Solutions Network, 2021).
25. T. Kukla *et al*., The resilience of Amazon tree cover to past and present drying. *Global Planet. Change* **202**, 103520 (2021). doi: 10.1016/j.gloplacha.2021.103520
26. Y. Qin *et al*., Carbon loss from forest degradation exceeds that from deforestation in the Brazilian Amazon. *Nat. Clim. Chang.* **11**, 442–448 (2021). doi: 10.1038/s41558-021-01026-5
27. J. Barichivich *et al*., Recent intensification of Amazon flooding extremes driven by strengthened Walker circulation. *Sci. Adv.* **4**, eaat8785 (2018). doi: 10.1126/sciadv.aat8785; pmid: 30255149
28. D. J. BertassoliJr *et al*., Spatiotemporal variations of riverine discharge within the Amazon basin during the late Holocene coincide with extratropical temperature anomalies. *Geophys. Res. Lett.* **46**, 9013–9022 (2019). doi: 10.1029/2019GL082936
29. I. S. A. A. Bezerra *et al*., Incision and aggradation phases of the Amazon River in central-eastern Amazonia during the late Neogene and Quaternary. *Geomorphology* **399**, 108073 (2022). doi: 10.1016/j.geomorph.2021.108073
30. H. Sato *et al*., Dry corridors opened by fire and low CO_2 in Amazonian rainforest during the Last Glacial Maximum. *Nat. Geosci.* **14**, 578–585 (2021). doi: 10.1038/s41561-021-00777-2
31. M. B. Bush *et al*., Widespread reforestation before European influence on Amazonia. *Science* **372**, 484–487 (2021). doi: 10.1126/science.abf3870; pmid: 33926948
32. O. Mottl *et al*., Global acceleration in rates of vegetation change over the past 18,000 years. *Science* **372**, 860–864 (2021). doi: 10.1126/science.abg1685; pmid: 34016781
33. S. Payette, in *Ecosystem Collapse and Climate Change*, J. G. Canadell, R. B. Jackson, Eds. (Springer, 2021), vol. 241, pp. 101–129; https://link.springer.com/chapter/10.1007/978-3-030-71330-0_5.
34. J. S. Albert, V. A. Tagliacollo, F. Dagosta, Diversification of Neotropical Freshwater Fishes. *Annu. Rev. Ecol. Evol. Syst.* **51**, 27–53 (2020). doi: 10.1146/annurev-ecolsys-011620-031032
35. N. Wunderling, J. F. Donges, J. Kurths, R. Winkelmann, Interacting tipping elements increase risk of climate domino effects under global warming. *Earth Syst. Dyn.* **12**, 601–619 (2021). doi: 10.5194/esd-12-601-2021
36. A. Cardil *et al*., Recent deforestation drove the spike in Amazonian fires. *Environ. Res. Lett.* **15**, 121003 (2020). doi: 10.1088/1748-9326/abcac7
37. J. C. O'Connor *et al*., Forests buffer against variations in precipitation. *Glob. Chang. Biol.* **27**, 4686–4696 (2021). doi: 10.1111/gcb.15763; pmid: 34319636
38. F. Hofhansl *et al*., Climatic and edaphic controls over tropical forest diversity and vegetation carbon storage. *Sci. Rep.* **10**, 5066 (2020). doi: 10.1038/s41598-020-61868-5; pmid: 32193471
39. A. Staal *et al*., Hysteresis of tropical forests in the 21st century. *Nat. Commun.* **11**, 4978 (2020). doi: 10.1038/s41467-020-18728-7; pmid: 33020475
40. G. S. Cooper, S. Willcock, J. A. Dearing, Regime shifts occur disproportionately faster in larger ecosystems. *Nat. Commun.* **11**, 1175 (2020). doi: 10.1038/s41467-020-15029-x; pmid: 32157098
41. B. M. Flores, M. Holmgren, White-sand savannas expand at the core of the Amazon after forest wildfires. *Ecosystems* **24**, 1624–1637 (2021). doi: 10.1007/s10021-021-00607-x
42. M. Goulding *et al*., Ecosystem-based management of Amazon fisheries and wetlands. *Fish Fish.* **20**, 138–158 (2019). doi: 10.1111/faf.12328
43. A. L. B. Fares, F. A. S. Nonato, T. S. Michelan, New records of the invasive macrophyte, *Urochloa arrecta* extend its range to eastern Brazilian Amazon altered freshwater ecosystems. *Acta Amazon.* **50**, 133–137 (2020). doi: 10.1590/1809-4392201903831
44. A. Rosell-Melé *et al*., Oil pollution in soils and sediments from the Northern Peruvian Amazon. *Sci. Total Environ.* **610-611**, 1010–1019 (2018). doi: 10.1016/j.scitotenv.2017.07.208; pmid: 28847095
45. Y. le Polain de Waroux *et al*., The restructuring of South American soy and beef production and trade under changing environmental regulations. *World Dev.* **121**, 188–202 (2019). doi: 10.1016/j.worlddev.2017.05.034
46. L. Ferrante, P. M. Fearnside, Countries should boycott Brazil over export-driven deforestation. *Nature* **601**, 318–318 (2022). doi: 10.1038/d41586-022-00094-7; pmid: 35042995
47. E. L. Bullock, C. E. Woodcock, C. SouzaJr.., P. Olofsson, Satellite-based estimates reveal widespread forest degradation in the Amazon. *Glob. Chang. Biol.* **26**, 2956–2969 (2020). doi: 10.1111/gcb.15029; pmid: 32022338
48. S. Hecht *et al*., in *Amazon Assessment Report 2021*, C. Nobre *et al.*, Eds. (United Nations Sustainable Development Solutions Network, 2021).
49. A. Franco-Solís, C. Montanía, Dynamics of deforestation worldwide: A structural decomposition analysis of agricultural land use in South America. *Land Use Policy* **109**, 105619 (2021). doi: 10.1016/j.landusepol.2021.105619
50. X. He, G. M. DePaula, W. Zhang, in *2021 Agricultural & Applied Economics Association Annual Meeting* (Agricultural & Applied Economics Association, 2021); https://ageconsearch.umn.edu/record/312818.
51. G. M. DePaula, L. Justino, in *2020 Annual Meeting* (Agricultural and Applied Economics Association, Kansas City, Missouri, 2020); https://ageconsearch.umn.edu/record/304482.
52. X. P. Song *et al*., Massive soybean expansion in South America since 2000 and implications for conservation. *Nat. Sustain.* **2021**, 784–792 (2021). doi: 10.1038/s41893-021-00729-z; pmid: 34377840
53. R. Rajão *et al*., The rotten apples of Brazil's agribusiness. *Science* **369**, 246–248 (2020). doi: 10.1126/science.aba6646; pmid: 32675358
54. R. T. Walker *et al*., Avoiding Amazonian catastrophes: Prospects for conservation in the 21st century. *One Earth* **1**, 202–215 (2019). doi: 10.1016/j.oneear.2019.09.009
55. B. A. Roy *et al*., New mining concessions could severely decrease biodiversity and ecosystem services in Ecuador. *Trop. Conserv. Sci.* 10.1177/1940082918780427 (2018). doi: 10.1177/1940082918780427
56. R. Andrade, in *The Political Economy of Hydropower in Southwest China and Beyond*, H.-S. S. Rousseau, Ed., International Political Economy Series (Palgrave Macmillan, 2021), pp. 275–293; https://link.springer.com/chapter/10.1007/978-3-030-59361-2_14.
57. V. S. Daga *et al*., Water diversion in Brazil threatens biodiversity. *Ambio* **49**, 165–172 (2020). doi: 10.1007/s13280-019-01189-8; pmid: 31030418
58. D. M. Larrea-Alcázara, N. Cuvi, J. F. Valentim, L. Diaz, S. Vidal, G. Palacio, in *Amazon Assessment Report 2021*, C. Nobre *et al.*, Eds. (United Nations Sustainable Development Solutions Network, 2021).
59. R. Heilmayr, L. L. Rausch, J. Munger, H. K. Gibbs, Brazil's Amazon Soy Moratorium reduced deforestation. *Nat. Food* **1**, 801–810 (2020). doi: 10.1038/s43016-020-00194-5
60. K. F. de Moraes *et al*., Climate change and bird extinctions in the Amazon. *PLOS ONE* **15**, e0236103 (2020). doi: 10.1371/journal.pone.0236103; pmid: 32678834
61. G. A. Herrera-R *et al*., The combined effects of climate change and river fragmentation on the distribution of Andean Amazon fishes. *Glob. Chang. Biol.* **26**, 5509–5523 (2020). doi: 10.1111/gcb.15285; pmid: 32785968
62. P. G. Zaninelli, C. G. Menéndez, M. Falco, N. López-Franca, A. F. Carril, Future hydroclimatological changes in South America based on an ensemble of regional climate models. *Clim. Dyn.* **52**, 819–830 (2019). doi: 10.1007/s00382-018-4225-0
63. M. Iturbide *et al*., An update of IPCC climate reference regions for subcontinental analysis of climate model data: Definition and aggregated datasets. *Earth Syst. Sci. Data* **12**, 2959–2970 (2020). doi: 10.5194/essd-12-2959-2020
64. J. C. Espinoza, J. A. Marengo, J. Schongart, J. C. Jimenez, The new historical flood of 2021 in the Amazon River compared to major floods of the 21st century: Atmospheric features in the context of the intensification of floods. *Weather Clim. Extrem.* **35**, 100406 (2022). doi: 10.1016/j.wace.2021.100406
65. D. A. Edmonds, R. L. Caldwell, E. S. Brondizio, S. M. O. Siani, Coastal flooding will disproportionately impact people on river deltas. *Nat. Commun.* **11**, 4741 (2020). doi: 10.1038/s41467-020-18531-4; pmid: 32994404
66. M. Moraes R., *et al.*, in *Amazon Assessment Report 2021*, C. Nobre *et al.*, Eds. (United Nations Sustainable Development Solutions Network, 2021).
67. V. Rull, A. C. Carnaval, *Neotropical Diversification: Patterns and Processes*, Fascinating Life Sciences (Springer International Publishing, 2020); http://link.springer.com/10.1007/978-3-030-31167-4.
68. S. Sanderson *et al*., The pace of modern life, revisited. *Mol. Ecol.* **31**, 1028–1043 (2022). doi: 10.1111/mec.16299; pmid: 34902193
69. D. R. Williams *et al*., Proactive conservation to prevent habitat losses to agricultural expansion. *Nat. Sustain.* **4**, 314–322 (2020). doi: 10.1038/s41893-020-00656-5
70. R. S. Nerem *et al*., Climate-change-driven accelerated sea-level rise detected in the altimeter era. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2022–2025 (2018). doi: 10.1073/pnas.1717312115; pmid: 29440401
71. A. J. Chadwick, M. P. Lamb, V. Ganti, Accelerated river avulsion frequency on lowland deltas due to sea-level rise. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 17584–17590 (2020). doi: 10.1073/pnas.1912351117; pmid: 32661152
72. C. H. L. Silva Junior *et al*., The Brazilian Amazon deforestation rate in 2020 is the greatest of the decade. *Nat. Ecol. Evol.* **5**, 144–145 (2021). doi: 10.1038/s41559-020-01368-x; pmid: 33349655
73. E. A. T. Matricardi *et al*., Long-term forest degradation surpasses deforestation in the Brazilian Amazon. *Science* **369**, 1378–1382 (2020). doi: 10.1126/science.abb3021; pmid: 32913104
74. M. Davis, S. Faurby, J. C. Svenning, Mammal diversity will take millions of years to recover from the current biodiversity crisis. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 11262–11267 (2018). doi: 10.1073/pnas.1804906115; pmid: 30322924
75. A. Esquivel-Muelbert *et al*., Compositional response of Amazon forests to climate change. *Glob. Chang. Biol.* **25**, 39–56 (2019). doi: 10.1111/gcb.14413; pmid: 30406962
76. A. Cendrero, L. M. Forte, J. Remondo, J. A. Cuesta-Albertos, Anthropocene geomorphic change. Climate or human activities? *Earths Futur.* **8**, e2019EF001305 (2020). doi: 10.1029/2019EF001305

77. G. H. E. Lense, J. C. Avanzi, T. C. Parreiras, R. L. Mincato, Effects of deforestation on water erosion rates in the Amazon region. *Agraria* **15**, 1–7 (2020). doi: 10.5039/agraria.v15i4a8500
78. C. W. Arnscheidt, D. H. Rothman, The balance of nature: A global marine perspective. *Ann. Rev. Mar. Sci.* **14**, 49–73 (2022). doi: 10.1146/annurev-marine-010318-095212; pmid: 34115541
79. V. Brovkin *et al*., Past abrupt changes, tipping points and cascading impacts in the Earth system. *Nat. Geosci.* **14**, 550–558 (2021). doi: 10.1038/s41561-021-00790-5
80. R. Hannah, R. Max, Forests and Deforestation—Our World in Data (2021); https://ourworldindata.org/forests-and-deforestation.
81. P. Borrelli *et al*., Land use and climate change impacts on global soil erosion by water (2015-2070). *Proc. Natl. Acad. Sci. U.S.A.* **117**, 21994–22001 (2020). doi: 10.1073/pnas.2001403117; pmid: 32839306
82. S. Jasechko, D. Perrone, Global groundwater wells at risk of running dry. *Science* **372**, 418–421 (2021). doi: 10.1126/science.abc2755; pmid: 33888642
83. P. Friedlingstein *et al*., Global Carbon Budget 2021. *Earth Syst. Sci. Data Discuss.* **2021**, 1–191 (2021).
84. C. Burton *et al*., South American fires and their impacts on ecosystems increase with continued emissions. *Clim. Resil. Sustain.* **1**, e8 (2022). doi: 10.1002/cli2.8
85. L. V. Gatti *et al*., Amazonia as a carbon source linked to deforestation and climate change. *Nature* **595**, 388–393 (2021). doi: 10.1038/s41586-021-03629-6; pmid: 34262208
86. R. Hamilton *et al*., Non-uniform tropical forest responses to the 'Columbian Exchange' in the Neotropics and Asia-Pacific. *Nat. Ecol. Evol.* **5**, 1174–1184 (2021). doi: 10.1038/s41559-021-01474-4; pmid: 34112995
87. G. L. Foster, D. L. Royer, D. J. Lunt, Future climate forcing potentially without precedent in the last 420 million years. *Nat. Commun.* **8**, 14845–14848 (2017). doi: 10.1038/ncomms14845; pmid: 28375201
88. P. Voosen, Ice shelf holding back keystone Antarctic glacier within years of failure. *Science* **374**, 1420–1421 (2021). doi: 10.1126/science.acz9833; pmid: 34914505
89. E. C. Pettit *et al*., in *AGU Fall Meeting* (AGU, 2021); https://agu.confex.com/agu/fm21/meetingapp.cgi/Paper/978762.
90. J. E. Tierney *et al*., Past climates inform our future. *Science* **370**, eaay3701 (2020). doi: 10.1126/science.aay3701; pmid: 33154110
91. S. Dangendorf *et al*., Persistent acceleration in global sea-level rise since the 1960s. *Nat. Clim. Chang.* **9**, 705–710 (2019). doi: 10.1038/s41558-019-0531-8
92. G. N. Inglis *et al*., Global mean surface temperature and climate sensitivity of the early Eocene Climatic Optimum (EECO), Paleocene-Eocene Thermal Maximum (PETM), and latest Paleocene. *Clim. Past* **16**, 1953–1968 (2020). doi: 10.5194/cp-16-1953-2020
93. L. L. Haynes, B. Hönisch, The seawater carbon inventory at the Paleocene-Eocene Thermal Maximum. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 24088–24095 (2020). doi: 10.1073/pnas.2003197117; pmid: 32929018
94. R. Guoyu *et al*., Characteristics, drivers and feedbacks of paleo-climatic variations and the implications for modern climate change research. *J. Quat. Sci.* **41**, 824–841 (2021).
95. C. Folke *et al*., Our future in the Anthropocene biosphere. *Ambio* **50**, 834–869 (2021). doi: 10.1007/s13280-021-01544-8; pmid: 33715097
96. D. Leclère *et al*., Bending the curve of terrestrial biodiversity needs an integrated strategy. *Nature* **585**, 551–556 (2020). doi: 10.1038/s41586-020-2705-y; pmid: 32908312
97. T. A. Gardner *et al*., Transparency and sustainability in global commodity supply chains. *World Dev.* **121**, 163–177 (2019). doi: 10.1016/j.worlddev.2018.05.025; pmid: 31481824
98. G. Chapron, Y. Epstein, J. V. López-Bao, A rights revolution for nature. *Science* **363**, 1392–1393 (2019). doi: 10.1126/science.aav5601; pmid: 30872530
99. R. Abramovay *et al*., in *Amazon Assessment Report 2021*, C. Nobre *et al*., Eds. (United Nations Sustainable Development Solutions Network, 2021).
100. C. Kremen, A. M. Merenlender, Landscapes that work for biodiversity and people. *Science* **362**, eaau6020 (2018). doi: 10.1126/science.aau6020; pmid: 30337381
101. N. E. Nedzel, *The Rule of Law, Economic Development, and Corporate Governance* (Edward Elgar Publishing, 2020).
102. A. Di Sacco *et al*., Ten golden rules for reforestation to optimize carbon sequestration, biodiversity recovery and livelihood benefits. *Glob. Chang. Biol.* **27**, 1328–1348 (2021). doi: 10.1111/gcb.15498; pmid: 33494123
103. D. Wilkinson, in *Locality and Identity: Environmental Issues in Law and Society* (Taylor and Francis, 2019), pp. 17–50; https://www.taylorfrancis.com/chapters/edit/10.4324/9780429449925-2/using-environmental-ethics-create-ecological-law-david-wilkinson.
104. B. E. Rollin, in *Problems of International Justice* (Taylor and Francis, 2019), pp. 124–1437; https://www.taylorfrancis.com/chapters/edit/10.4324/9780429303111-8/environmental-ethics-international-justice-bernard-rollin.
105. J. M. Guayasamin *et al*., Biodiversity conservation: local and global consequences of the application of 'rights of nature' by Ecuador. *Neotrop. Biodivers.* **7**, 541–545 (2021).
106. C. D. Stone, *Should Trees Have Standing? Law, Morality, and the Environment | Environment & Society Portal* (Oxford Univ. Press, ed. 3, 2010); https://www.environmentandsociety.org/mml/should-trees-have-standing-law-morality-and-environment.
107. D. R. Boyd, *Rights of Nature: A Legal Revolution That Could Save the World* (ECW Press, 2017).
108. E. L. O'Donnell, J. Talbot-Jones, Creating legal rights for rivers: Lessons from Australia, New Zealand, and India. *Ecol. Soc.* **23**, art7 (2018). doi: 10.5751/ES-09854-230107
109. X. Feng *et al*., How deregulation, drought and increasing fire impact Amazonian biodiversity. *Nature* **597**, 516–521 (2021). doi: 10.1038/s41586-021-03876-7; pmid: 34471291
110. P. Higgins, D. Short, N. South, Protecting the planet: A proposal for a law of ecocide. *Crime Law Soc. Change* **59**, 251–266 (2013). doi: 10.1007/s10611-013-9413-6
111. M. Lynch, Regressive prosecutors: Law and order politics and practices in Trump's DOJ. *Hastings J. Crime Punish* **1**, 195 (2020).
112. T. A. P. West, P. M. Fearnside, Brazil's conservation reform and the reduction of deforestation in Amazonia. *Land Use Policy* **100**, 105072 (2021). doi: 10.1016/j.landusepol.2020.105072
113. S. Díaz *et al*., Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science* **366**, eaax3100 (2019). doi: 10.1126/science.aax3100; pmid: 31831642
114. E. S. Brondízio *et al*., Locally based, regionally manifested, and globally relevant: Indigenous and local knowledge, values, and practices for nature. *Annu. Rev. Environ. Resour.* **46**, 481–509 (2021). doi: 10.1146/annurev-environ-012220-012127
115. D. E. Bunker *et al*., Species loss and aboveground carbon storage in a tropical forest. *Science* **310**, 1029–1031 (2005). doi: 10.1126/science.1117682; pmid: 16239439
116. C. E. Wheeler *et al*., Carbon sequestration and biodiversity following 18 years of active tropical forest restoration. *For. Ecol. Manage.* **373**, 44–55 (2016). doi: 10.1016/j.foreco.2016.04.025
117. B. Soares-Filho, R. Rajão, Traditional conservation strategies still the best option. *Nat. Sustain.* **1**, 608–610 (2018). doi: 10.1038/s41893-018-0179-9
118. R. Agudelo *et al*., Land use planning in the Amazon basin: Challenges from resilience thinking. *Ecol. Soc.* **25**, 18 (2020).
119. INPE — Coordenação-Geral de Observação da Terra, "PRODES 2020: Monitoramento do Desmatamento da Floresta Amazônica Brasileira por Satélite" (2020); http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes.
120. N. Stern, J. Stiglitz, C. Taylor, The economics of immense risk, urgent action and radical change: Towards new approaches to the economics of climate change. *J. Econ. Methodol.* **29**, 181–216 (2022). doi: 10.1080/1350178X.2022.2040740
121. A. A. Min-Venditti, G. W. Moore, F. Fleischman, What policies improve forest cover? A systematic review of research from Mesoamerica. *Glob. Environ. Change* **47**, 21–27 (2017). doi: 10.1016/j.gloenvcha.2017.08.010
122. R. Hajjar, J. A. Oldekop, Research frontiers in community forest management. *Curr. Opin. Environ. Sustain.* **32**, 119–125 (2018). doi: 10.1016/j.cosust.2018.06.003
123. R. Brouwer, R. Pinto, A. Dugstad, S. Navrud, The economic value of the Brazilian Amazon rainforest ecosystem services: A meta-analysis of the Brazilian literature. *PLOS ONE* **17**, e0268425 (2022). doi: 10.1371/journal.pone.0268425; pmid: 35588116
124. M. Paldam, *The Grand Pattern of Development and the Transition of Institutions* (Cambridge Univ. Press, 2021).
125. IRENA, "World Energy Transitions Outlook: 1.5°C Pathway" (Abu Dhabi, 2021); https://irena.org/publications/2021/Jun/World-Energy-Transitions-Outlook.
126. C. Wilson *et al*., Granular technologies to accelerate decarbonization. *Science* **368**, 36–39 (2020). doi: 10.1126/science.aaz8060; pmid: 32241941
127. D. Welsby, J. Price, S. Pye, P. Ekins, Unextractable fossil fuels in a 1.5 °C world. *Nature* **597**, 230–234 (2021). doi: 10.1038/s41586-021-03821-8; pmid: 34497394
128. J. Rockström, O. Edenhofer, J. Gaertner, F. DeClerck, Planet-proofing the global food system. *Nat. Food* **1**, 3–5 (2020). doi: 10.1038/s43016-019-0010-4
129. A. A. Zuniga-Teran *et al*., Challenges of mainstreaming green infrastructure in built environment professions. *J. Environ. Plann. Manage.* **63**, 710–732 (2019). doi: 10.1080/09640568.2019.1605890
130. C. Levis *et al*., Help restore Brazil's governance of globally important ecosystem services. *Nat. Ecol. Evol.* **4**, 172–173 (2020). doi: 10.1038/s41559-019-1093-x; pmid: 32015426
131. S. Athayde *et al*., in *Amazon Assessment Report 2021*, C. Nobre *et al*., Eds. (United Nations Sustainable Development Solutions Network, 2021).
132. IPBES, (2019), *Summary for Policymakers of the Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services* (IPBES secretariat, 2019); https://www.ipbes.net/global-assessment.
133. M. H. Ruckelshaus *et al*., The IPBES Global Assessment: Pathways to Action. *Trends Ecol. Evol.* **35**, 407–414 (2020). doi: 10.1016/j.tree.2020.01.009; pmid: 32294422
134. P. R. Shukla *et al*., "Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems" (IPCC, 2019).
135. K. O. Winemiller *et al*., Balancing hydropower and biodiversity in the Amazon, Congo, and Mekong. *Science* **351**, 128–129 (2016). doi: 10.1126/science.aac7082; pmid: 26744397
136. B. R. Scanlon, I. Jolly, M. Sophocleous, L. Zhang, Global impacts of conversions from natural to agricultural ecosystems on water resources: Quantity versus quality. *Water Resour. Res.* **43**, 3437 (2007). doi: 10.1029/2006WR005486
137. E. C. Ellis, K. K. Goldewijk, S. Siebert, D. Lightman, N. Ramankutty, Anthropogenic transformation of the biomes, 1700 to 2000. *Glob. Ecol. Biogeogr.* **19**, 589–606 (2010). doi: 10.1111/j.1466-8238.2010.00540.x
138. J. F. Bastin *et al*., The global tree restoration potential. *Science* **365**, 76–79 (2019). doi: 10.1126/science.aax0848; pmid: 31273120
139. D. C. da Cruz, J. M. R. Benayas, G. C. Ferreira, S. R. Santos, G. Schwartz, An overview of forest loss and restoration in the Brazilian Amazon. *New For.* **52**, 1–16 (2021). doi: 10.1007/s11056-020-09777-3
140. N. Saintilan *et al*., Thresholds of mangrove survival under rapid sea level rise. *Science* **368**, 1118–1121 (2020). doi: 10.1126/science.aba2656; pmid: 32499441
141. C. D. Storlazzi *et al*., Most atolls will be uninhabitable by the mid-21st century because of sea-level rise exacerbating wave-driven flooding. *Sci. Adv.* **4**, eaap9741 (2018). doi: 10.1126/sciadv.aap9741; pmid: 29707635
142. T. E. Törnqvist, K. L. Jankowski, Y.-X. Li, J. L. González, Tipping points of Mississippi Delta marshes due to accelerated sea-level rise. *Sci. Adv.* **6**, eaaz5512 (2020). doi: 10.1126/sciadv.aaz5512; pmid: 32494741
143. J. S. Albert, J. M. Craig, V. A. Tagliacollo, P. Petry, in *Mountains, Climate, and Biodiversity*, C. Hoorn, A. Perrigo, A. Antonelli, Eds. (John Wiley, ed. 1, 2018), pp. 273–294.
144. L. W. Alvarez, Mass extinctions caused by large bolide impacts. *Phys. Today* **40**, 24–33 (1987). doi: 10.1063/1.881078; pmid: 11542219
145. B. Gutenberg, C. F. Richter, *Seismicity of the Earth and Associated Phenomena* (Princeton Univ. Press, ed. 2, 1954).
146. C. G. A. Harrison, Rates of continental erosion and mountain building. *Geol. Rundsch.* **83**, 431–447 (1994). doi: 10.1007/BF00210556
147. N. E. Matthews, J. A. Vazquez, A. T. Calvert, Age of the Lava Creek supereruption and magma chamber assembly at Yellowstone based on 40Ar/39Ar and U-Pb dating of sanidine and zircon crystals. *Geochem. Geophys. Geosyst.* **16**, 2508–2528 (2015). doi: 10.1002/2015GC005881

148. P. Papale, W. Marzocchi, Volcanic threats to global society. *Science* **363**, 1275–1276 (2019). doi: 10.1126/science.aaw7201; pmid: 30898915
149. Y. Sawai, Y. Namegaya, Y. Okamura, K. Satake, M. Shishikura, Challenges of anticipating the 2011 Tohoku earthquake and tsunami using coastal geology. *Geophys. Res. Lett.* **39**, L21309 (2012). doi: 10.1029/2012GL053692
150. C. Wang, R. N. Mitchell, J. B. Murphy, P. Peng, C. J. Spencer, The role of megacontinents in the supercontinent cycle. *Geology* **49**, 402–406 (2021). doi: 10.1130/G47988.1
151. C. R. Neal, J. J. Mahoney, L. W. Kroenke, R. A. Duncan, M. G. Petterson, The Ontong Java Plateau. *Geophys. Monogr.* **100**, 183–216 (1997).
152. P. A. Bland, N. A. Artemieva, The rate of small impacts on Earth. *Meteorit. Planet. Sci.* **41**, 607–631 (2006). doi: 10.1111/j.1945-5100.2006.tb00485.x
153. W. C. Clark, Scales of climate impacts. *Clim. Change* **7**, 5–27 (1985). doi: 10.1007/BF00139438
154. M. Nearing, F. F. Pruski, M. R. O'Neal, Expected climate change impacts on soil erosion rates: A review. *J. Soil Water Conserv.* **59**, 43–50 (2004).
155. M. E. Clapham, P. R. Renne, Flood basalts and mass extinctions. *Annu. Rev. Earth Planet. Sci.* **47**, 275–303 (2019). doi: 10.1146/annurev-earth-053018-060136
156. P. D. Gingerich, Rates of evolution. *Annu. Rev. Ecol. Evol. Syst.* **40**, 657–675 (2009). doi: 10.1146/annurev.ecolsys.39.110707.173457
157. C. A. Suarez, M. Edmonds, A. P. Jones, Earth catastrophes and their impact on the carbon cycle. *Elements* **15**, 301–306 (2019). doi: 10.2138/gselements.15.5.301
158. S. Siebert *et al*., A global data set of the extent of irrigated land from 1900 to 2005. *Hydrol. Earth Syst. Sci.* **19**, 1521–1545 (2015). doi: 10.5194/hess-19-1521-2015
159. S. Hu, Z. Niu, Y. Chen, L. Li, H. Zhang, Global wetlands: Potential distribution, wetland loss, and status. *Sci. Total Environ.* **586**, 319–327 (2017). doi: 10.1016/j.scitotenv.2017.02.001; pmid: 28190574
160. J. D. Milliman, K. L. Farnsworth, *River Discharge to the Coastal Ocean: A Global Synthesis* (Cambridge Univ. Press, 2011).
161. M. T. H. van Vliet *et al*., Global river discharge and water temperature under climate change. *Glob. Environ. Change* **23**, 450–464 (2013). doi: 10.1016/j.gloenvcha.2012.11.002
162. K. C. Seto, M. Fragkias, B. Güneralp, M. K. Reilly, A meta-analysis of global urban land expansion. *PLOS ONE* **6**, e23777 (2011). doi: 10.1371/journal.pone.0023777; pmid: 21876770
163. M. C. Hansen *et al*., High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013). doi: 10.1126/science.1244693; pmid: 24233722
164. P. Val *et al*., in *Amazon Assessment Report 2021*, C. Nobre *et al*., Eds. (United Nations Sustainable Development Solutions Network, 2021).
165. L. Yahdjian, O. E. Sala, Climate change impacts on South American rangelands. *Rangelands* **30**, 34–39 (2008). doi: 10.2111/1551-501X(2008)30[34:CCIOSA]2.0.CO;2
166. M. Finer, C. N. Jenkins, S. L. Pimm, B. Keane, C. Ross, Oil and gas projects in the Western Amazon: Threats to wilderness, biodiversity, and indigenous peoples. *PLOS ONE* **3**, e2932 (2008). doi: 10.1371/journal.pone.0002932; pmid: 18716679
167. L. E. O. C. Aragão *et al*., 21st Century drought-related fires counteract the decline of Amazon deforestation carbon emissions. *Nat. Commun.* **9**, 536–12 (2018). doi: 10.1038/s41467-017-02771-y; pmid: 29440640
168. M. V. F. Silveira *et al*., Drivers of fire anomalies in the Brazilian Amazon: Lessons learned from the 2019 fire crisis. *Land (Basel)* **9**, 516 (2020). doi: 10.3390/land9120516
169. A. Ivanov *et al*., Siberian Traps large igneous province: Evidence for two flood basalt pulses around the Permo-Triassic boundary and in the Middle Triassic, and contemporaneous granitic magmatism. *Earth Sci. Rev.* **122**, 58–76 (2013). doi: 10.1016/j.earscirev.2013.04.001
170. R. N. Singh, K. R. Gupta, Workshop yields new insight into volcanism at Deccan Traps, India. *Eos* **75**, 356–356 (1994). doi: 10.1029/94EO01005
171. J. S. Albert, P. Val, C. Hoorn, The changing course of the Amazon River in the Neogene: Center stage for Neotropical diversification. *Neotrop. Ichthyol.* **16**, e180033 (2018). doi: 10.1590/1982-0224-20180033
172. N. Espurt *et al*., Flat subduction dynamics and deformation of the South American plate: Insights from analog modeling. *Tectonics* **27**, TC3011 (2008). doi: 10.1029/2007TC002175
173. A. Rocha *et al*., Pleistocene climatic oscillations associated with landscape heterogeneity of the South American dry diagonal explains the phylogeographic structure of the narrow-billed woodcreeper (*Lepidocolaptes angustirostris*, Dendrocolaptidae). *J. Avian Biol.* **51**, jav.02537 (2020). doi: 10.1111/jav.02537
174. F. Wittmann, W. J. Junk, in *The Wetland Book*, Finlayson C., Milton G., Prentice R., Davidson N., Eds. (Springer, 2016), pp. 1–20; https://link.springer.com/referenceworkentry/10.1007/978-94-007-6173-5_83-2.

ACKNOWLEDGMENTS

The SPA acknowledges generous financial support from the Gordon and Betty Moore Foundation and the Charles Stewart Mott Foundation. We thank the following for discussions: A. Antonelli, W. G. R. Crampton, G. Destouni, A. E. Magurran, T. Oberdorff, R. E. Reis, K. O. Winemiller, and W. J. Ripple. **Funding**: This work was supported in part by US National Science Foundation (NSF) 0614334, 0741450, and 1354511 to J.S.A.; NSF 1745562 and 1926928 to A.C.C.; Universidad San Francisco de Quito HUBI 5466, 16871, and 16808 to J.M.G.; Swiss National Science Foundation P4P4PB-199187 to J.D.C.; Trond Mohn Stiftelse (TMS) and University of Bergen TMS2022STG03 to S.G.A.F.; Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) 2012/50260-6, and 2018/23899-2 to L.G.L.; Fundação de Amparo à Pesquisa do Estado do Amazonas and US Agency for International Development 311732/2020-8 to C.C.R.; and Brazilian National Council for Scientific and Technological Development (CNPq) to C.A.Q. and 310871/2017-4 to L.G.L. **Author contributions:** J.S.A., A.C.C., L.G.L., C.C.R., D.R., J.D.C., J.M.G., and C.U.U. compiled and synthesized the biological datasets; S.G.A.F., Y.F., J.J.P.F., C.H., G.H.d.M., C.A.Q., and P.V. compiled and synthesized the geological datasets. C.A.N., A.C.E., J.A., and N.N. coordinated SPA activities. J.S.A. wrote the preliminary draft of the manuscript, and all authors contributed to the final draft; J.S.A., P.V., and N.N. prepared the figures. **Competing interests:** None declared. **Data and materials availability:** All data come from previously published works that are referenced in the table and figure legends.

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

Submitted 7 February 2022; resubmitted 4 March 2022
Accepted 22 November 2022
10.1126/science.abo5003



AAAS.ORG/COMMUNITY







AAAS' Member Community is a one-stop destination for scientists and STEM enthusiasts alike. It's "Where Science Gets Social": a community where facts matter, ideas are big and there's always a reason to come hang out, share, discuss and explore.

Member COMMUNITY
AAAS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

REVIEW SUMMARY

TROPICAL FOREST

The drivers and impacts of Amazon forest degradation

David M. Lapola*, Patricia Pinho, Jos Barlow, Luiz E. O. C. Aragão, Erika Berenguer, Rachel Carmenta, Hannah M. Liddy, Hugo Seixas, Camila V. J. Silva, Celso H. L. Silva-Junior, Ane A. C. Alencar, Liana O. Anderson, Dolors Armenteras, Victor Brovkin, Kim Calders, Jeffrey Chambers, Louise Chini, Marcos H. Costa, Bruno L. Faria, Philip M. Fearnside, Joice Ferreira, Luciana Gatti, Victor Hugo Gutierrez-Velez, Zhangang Han, Kathleen Hibbard, Charles Koven, Peter Lawrence, Julia Pongratz, Bruno T. T. Portela, Mark Rounsevell, Alex C. Ruane, Rüdiger Schaldach, Sonaira S. da Silva, Celso von Randow, Wayne S. Walker

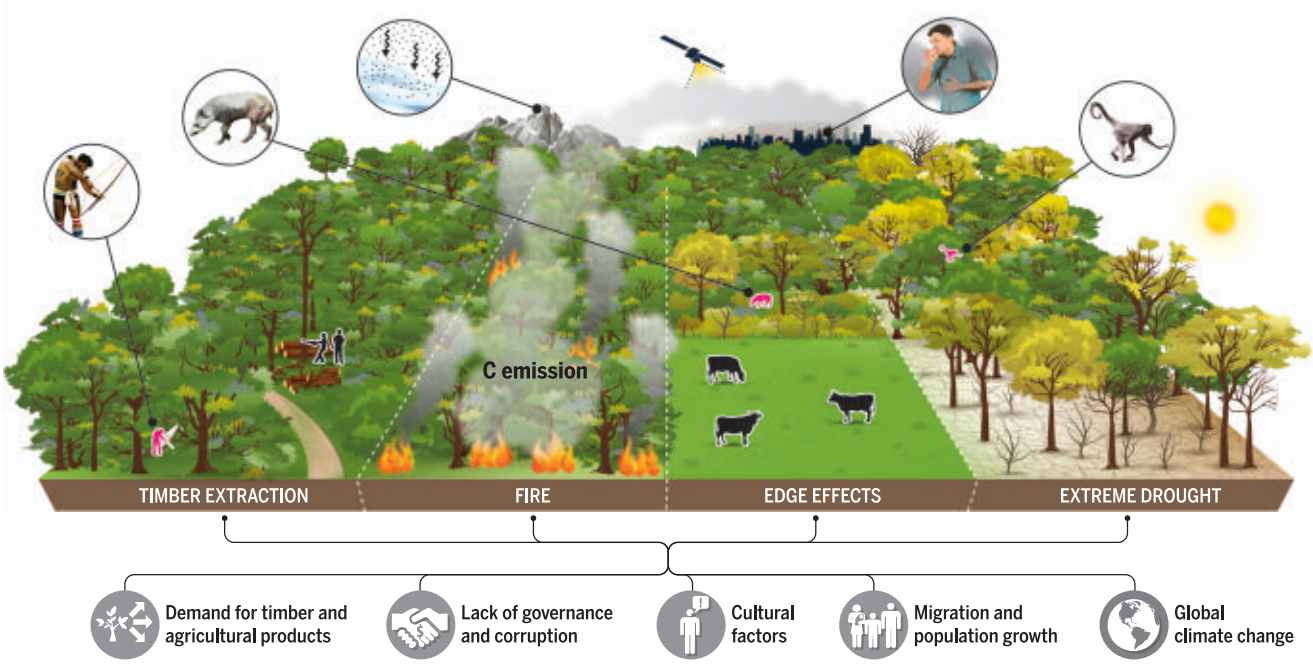
BACKGROUND: Most analyses of land-use and land-cover change in the Amazon forest have focused on the causes and effects of deforestation. However, anthropogenic disturbances cause degradation of the remaining Amazon forest and threaten their future. Among such disturbances, the most important are edge effects (due to deforestation and the resulting habitat fragmentation), timber extraction, fire, and extreme droughts that have been intensified by human-induced climate change. We synthesize knowledge on these disturbances that lead to Amazon forest degradation, including their causes and impacts, possible future extents, and some of the interventions required to curb them.

ADVANCES: Analysis of existing data on the extent of fire, edge effects, and timber extraction between 2001 and 2018 reveals that 0.36 $\times 10^6$ km^2 (5.5%) of the Amazon forest is under some form of degradation, which corresponds to 112% of the total area deforested in that period. Adding data on extreme droughts increases the estimate of total degraded area to 2.5 $\times 10^6$ km^2, or 38% of the remaining Amazonian forests. Estimated carbon loss from these forest disturbances ranges from 0.05 to 0.20 Pg C year^{-1} and is comparable to carbon loss from deforestation (0.06 to 0.21 Pg C year^{-1}). Disturbances can bring about as much biodiversity loss as deforestation itself, and forests degraded by fire and timber extraction can have a 2 to 34% reduction in dry-season evapotranspiration. The underlying drivers of disturbances (e.g., agricultural expansion or demand for timber) generate material benefits for a restricted group of regional and global actors, whereas the burdens permeate across a broad range of scales and social groups ranging from nearby forest dwellers to urban residents of Andean countries. First-order 2050 projections indicate that the four main disturbances will remain a major threat and source of carbon fluxes to the atmosphere, independent of deforestation trajectories.

OUTLOOK: Whereas some disturbances such as edge effects can be tackled by curbing deforestation, others, like constraining the increase in extreme droughts, require additional measures, including global efforts to reduce greenhouse gas emissions. Curbing degradation will also require engaging with the diverse set of actors that promote it, operationalizing effective monitoring of different disturbances, and refining policy frameworks such as REDD+. These will all be supported by rapid and multidisciplinary advances in our socioenvironmental understanding of tropical forest degradation, providing a robust platform on which to co-construct appropriate policies and programs to curb it. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: dmlapola@unicamp.br
Cite this article as D. M. Lapola *et al.*, *Science* **379**, eabp8622 (2023). DOI: 10.1126/science.abp8622

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abp8622



CREDIT: ALEX ARGOZINO/STUDIO ARGOZINO

An overview of tropical forest degradation processes in the Amazon. Underlying drivers (a few of which are shown in gray at the bottom) stimulate disturbances (timber extraction, fire, edge effects, and extreme drought) that cause forest degradation. A satellite illustrates the attempts to estimate degradation's spatial extent and associated carbon losses. Impacts (in red and insets) are either local—causing biodiversity losses or affecting forest-dweller livelihoods—or remote, for example, with smoke affecting people's health in cities or causing the melting of Andean glaciers owing to black carbon deposition.

REVIEW

TROPICAL FOREST

The drivers and impacts of Amazon forest degradation

David M. Lapola[1]*, Patricia Pinho[2], Jos Barlow[3], Luiz E. O. C. Aragão[4,5], Erika Berenguer[3,6], Rachel Carmenta[7], Hannah M. Liddy[8,9], Hugo Seixas[1], Camila V. J. Silva[2,3,10], Celso H. L. Silva-Junior[11,12,13], Ane A. C. Alencar[2], Liana O. Anderson[14], Dolors Armenteras[15], Victor Brovkin[16], Kim Calders[17,18], Jeffrey Chambers[19], Louise Chini[20], Marcos H. Costa[21], Bruno L. Faria[22], Philip M. Fearnside[23], Joice Ferreira[24], Luciana Gatti[4], Victor Hugo Gutierrez-Velez[25], Zhangang Han[26], Kathleen Hibbard[27], Charles Koven[19], Peter Lawrence[28], Julia Pongratz[16,29], Bruno T. T. Portela[23], Mark Rounsevell[30,31], Alex C. Ruane[9], Rüdiger Schaldach[32], Sonaira S. da Silva[33], Celso von Randow[4], Wayne S. Walker[34]

Approximately 2.5 × 10^6 square kilometers of the Amazon forest are currently degraded by fire, edge effects, timber extraction, and/or extreme drought, representing 38% of all remaining forests in the region. Carbon emissions from this degradation total up to 0.2 petagrams of carbon per year (Pg C $year^{-1}$), which is equivalent to, if not greater than, the emissions from Amazon deforestation (0.06 to 0.21 Pg C $year^{-1}$). Amazon forest degradation can reduce dry-season evapotranspiration by up to 34% and cause as much biodiversity loss as deforestation in human-modified landscapes, generating uneven socioeconomic burdens, mainly to forest dwellers. Projections indicate that degradation will remain a dominant source of carbon emissions independent of deforestation rates. Policies to tackle degradation should be integrated with efforts to curb deforestation and complemented with innovative measures addressing the disturbances that degrade the Amazon forest.

Tropical forests are critical for Earth's climate, biodiversity, local well-being and livelihoods, and humanity at large (*1*). They are also a hotspot for CO_2 emissions to the atmosphere, largely as a result of deforestation and other anthropogenic disturbances (*2*). Most analyses of land-use and land-cover changes in tropical forests have focused on the causes and effects of deforestation (*3*–*5*). However, other, less-well-studied anthropogenic disturbances also threaten the future of tropical forests. These disturbances include edge effects, selective logging, fire, and extreme drought, which have been intensified by human-induced climate change.

In the Amazon forest, the extent and long-term effects of such anthropogenic disturbances on the terrestrial carbon cycle, ecosystem functioning, and livelihoods of local populations are beginning to be understood and differentiated from deforestation impacts (*6*). These disturbances often co-occur and repeat multiple times and greatly increase the impact on forest condition and biodiversity (*7*). Many of the effects of these disturbances also occur over longer time scales. For instance, ongoing tree mortality after disturbance means that forests can continue to emit more carbon for decades after the disturbance (*8*, *9*), such that current estimates of the total carbon loss tied to degradation are comparable to, if not greater than, carbon loss from deforestation (*10*–*16*). Moreover, the reduced provision of ecosystem services resulting from such anthropogenic disturbances appears to disproportionately affect local livelihoods (*17*–*19*).

A recent study of the Amazon showed that only 14% of degraded forests were later deforested over a period of 22 years (*11*), suggesting that these are partially independent processes. Understanding and representing degradation as a process separate from deforestation is thus critical for improving observation networks, climate change and conservation policies, as well as modeling the resilience of the Amazon forest and its human populations, in light of ongoing land-cover and land-use changes and increased frequency of climate extremes.

In this Review, we (i) identify proximate and underlying drivers of disturbances related to pan-Amazon forest degradation; (ii) provide estimates of uncertainties in the total degraded forest area; (iii) assess the ecological impacts of degradation; and (iv) discuss the distribution of benefits and burdens among stakeholder groups. We then (v) examine our current ability to project Amazon degradation with existing data on disturbances and (vi) highlight the scientific advances required to understand and address forest degradation in Amazonia and other tropical forests.

Defining degradation and disturbance

Although many distinct definitions of forest degradation exist (*20*, *21*), for this Review we consider tropical forest degradation as a transitory or long-term (10^1- to 10^3-year time scale) deleterious change in forest condition. Condition includes functions, properties, or services such as, but not restricted to, carbon storage, biological productivity, species composition, forest structure, local atmospheric moisture, or uses and values of the forest to humans. Changes in forest condition can be determined through comparisons with a previous undisturbed baseline or inferred spatially using comparable undisturbed forests. Here, we focus on degradation driven by four human-induced disturbances (Fig. 1): extreme droughts, edge effects resulting from habitat fragmentation, timber extraction, and forest fires.

Extreme droughts have become increasingly frequent in the Amazon as land-use change and human-induced climate change progress (*22*), affecting tree mortality, fire incidence, and carbon emissions to the atmosphere (*23*–*25*). Deforestation leads to habitat fragmentation, including the edge, area, and isolation effects that are known drivers of changes in ecological condition (*12*, *26*). We focus mostly on edge effects, which are the changes in ecological and biophysical parameters that occur in forests adjacent to anthropogenic land uses (*9*). Timber extraction includes the legal and illegal selective logging that takes place in standing forests

[1]Laboratório de Ciência do Sistema Terrestre – LabTerra, Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura - CEPAGRI, Universidade Estadual de Campinas, Campinas, SP, Brazil. [2]Instituto de Pesquisas Ambientais da Amazônia, Brasília, DF, Brazil. [3]Lancaster Environment Centre, Lancaster University, Lancaster, UK. [4]Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP, Brazil. [5]Geography, University of Exeter, Exeter, UK. [6]Environmental Change Institute, University of Oxford, Oxford, UK. [7]University of East Anglia, Norwich, UK. [8]Columbia Climate School, Columbia University, New York, NY, USA. [9]NASA Goddard Institute for Space Studies, New York, NY, USA. [10]BeZero Carbon Ltd, London, UK. [11]Institute of Environment and Sustainability, University of California, Los Angeles, CA, USA. [12]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. [13]Programa de Pós-graduação em Biodiversidade e Conservação, Universidade Federal do Maranhão - UFMA, São Luís, MA, Brazil. [14]Centro Nacional de Monitoramento e Alertas de Desastres Naturais, São José dos Campos, SP, Brazil. [15]Universidad Nacional de Colombia, Bogotá, Colombia. [16]Max Planck Institute for Meteorology, Hamburg, Germany. [17]Computational & Applied Vegetation Ecology Laboratory, Department of Environment, Ghent University, Belgium. [18]School of Forest Sciences, University of Eastern Finland, Joensuu, Finland. [19]Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [20]University of Maryland, College Park, MD, USA. [21]Universidade Federal de Viçosa, Viçosa, MG, Brazil. [22]Instituto Federal de Educação, Ciência e Tecnologia do Norte de Minas Gerais, Diamantina, MG, Brazil. [23]Instituto Nacional de Pesquisas da Amazônia, Manaus, AM, Brazil. [24]Empresa Brasileira de Pesquisa Agropecuária, Belém, PA, Brazil. [25]Temple University, Philadelphia, PA, USA. [26]Beijing Normal University, Beijing, China. [27]National Aeronautics and Space Administration Headquarters, Washington, DC, USA. [28]National Center for Atmospheric Research, Boulder, CO, USA. [29]Ludwig-Maximilians University of Munich, Munich, Germany. [30]Karlsruhe Institute of Technology, Karlsruhe, Germany. [31]University of Edinburgh, Edinburgh, UK. [32]University of Kassel, Kassel, Germany. [33]Universidade Federal do Acre, Cruzeiro do Sul, AC, Brazil. [34]Woodwell Climate Research Center, Falmouth, MA, USA.
*Corresponding author. Email: dmlapola@unicamp.br

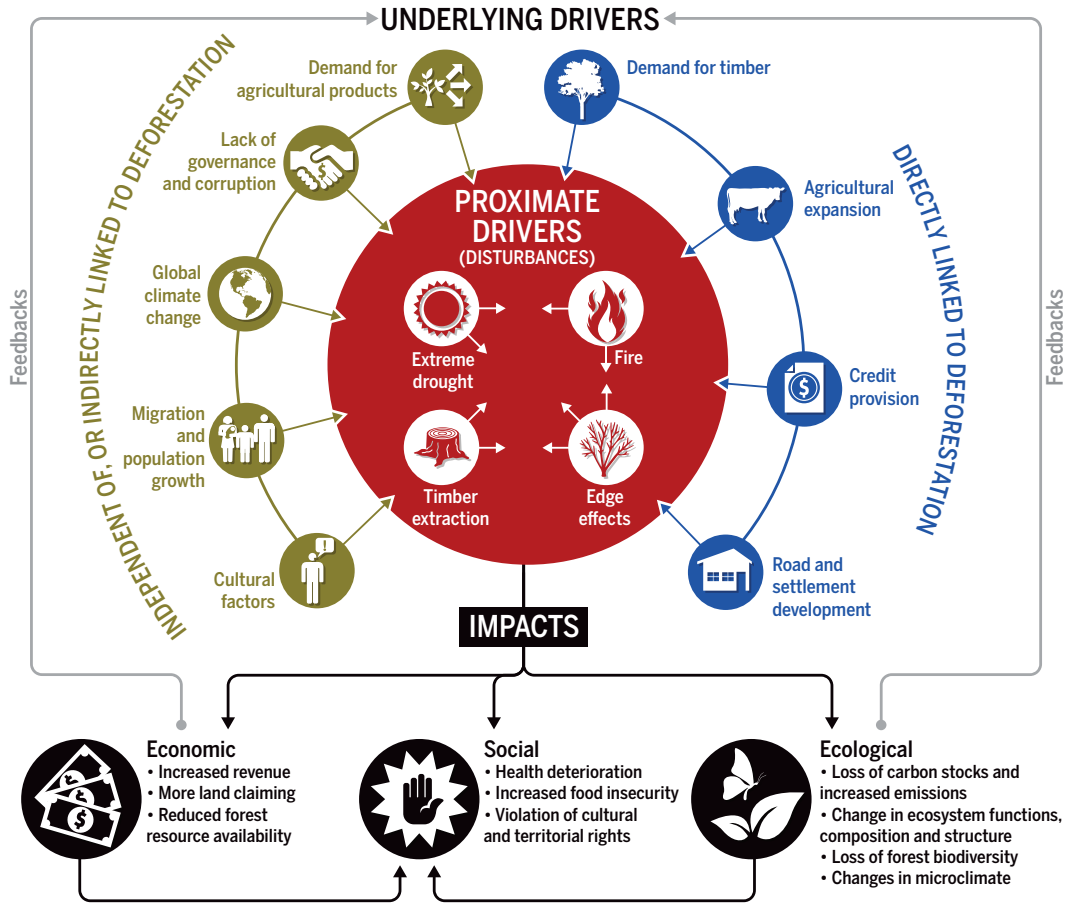


Fig. 1. A conceptual model of the drivers, impacts, and feedbacks of Amazon forest degradation. Some of the underlying drivers of disturbance (outer edge of the circle) are also directly related to deforestation processes (blue), whereas others are not (olive). The four main disturbances leading to degradation—extreme drought, fire, timber extraction, and edge effects—are intrinsically interrelated and can feed back to each other (see "Underlying drivers of disturbance" section). These disturbances cause economic, social, and ecological impacts that can be directly linked to each other (e.g., reduced forest resource availability leading to food insecurity). These impacts can also feed back to influence both the underlying and proximate drivers (respectively exemplified by revenue related to degradation causing local migratory movements and changes in microclimate causing increased fire incidence).

(*27*, *28*). Forest fires include all fires in standing forests (*29*); these cause degradation as Amazonian species have little or no evolutionary adaptations to fire. This list is not comprehensive; for example, heat stress, isolation effects, nontimber forest product extraction, and defaunation could all alter forest condition. However, the four disturbances that we focus on can all be studied across the Amazon by using available satellite data and image processing methods, and have the best-quantified links with forest structure and carbon stocks. We do not consider natural disturbances (e.g., blowdowns) to be degradation unless they interact with anthropogenic disturbances (*30*).

To evaluate degradation, it must be differentiated from deforestation. Conceptually, this is simple. Deforestation involves a change in land cover (e.g., loss of canopy cover to below a certain threshold) and generally a change in land use (e.g., from forest to agriculture or urban land use) (*31*). By contrast, whereas land use may or may not change during the process of degradation, land cover does not (i.e., forest remains forest). However, this conceptual clarity can break down when monitoring at scale. First, satellite-based monitoring of forests cannot easily discern changes in land use—areas affected by severe disturbances, such as thrice-burned forests, can be classified as deforested even though the land use has not changed. Second, some deliberate deforestation may be confused with degradation, with actors aiming to escape legal prosecution by using successive fires and other disturbances to gradually reduce tree cover over time. Although we do not attempt to integrate these nuances in this Review, future monitoring would benefit from considering four land-cover classes within the degradation-deforestation continuum: (i) undisturbed forest; (ii) degraded forest, where forest cover remains above a critical threshold and land-use change has not occurred; (iii) deforestation caused by successive or severe disturbances, where forest cover falls below a critical threshold of forest canopy structure but land-use change does not occur; and (iv) clear-cut deforestation, where forest cover falls below a critical threshold as a result of land-use change. Differentiating between the latter three classes is key for applying legal processes and can be supported by longer-term assessments, consideration of the geometric patterns of change (burned edges are rarely linear), and ground visits.

Underlying drivers of disturbance

The disturbances that cause degradation share a range of underlying drivers operating at the regional or landscape scale (e.g., lack of governance, presence of roads, or demand for local foods) or stemming from national or global influences (e.g., market demand for commodities, credit, and climate change) (Fig. 1). Many of these drivers are linked with deforestation (*5*). For example, agricultural expansion into forested lands increases the exposure of the

CREDIT: ALEX ARGOZINO/STUDIO ARGOZINO

remaining forests to edge effects, timber extraction, and the agricultural ignition sources that start many forest fires (*32*). Other key drivers of forest disturbance are, however, largely independent of the Amazonian deforestation process. Some timber extraction occurs in remote regions, far from the deforestation frontier; fires can extend deep into forested areas in drought years (*25*); and droughts are widespread across the basin (*22*, *33*).

The underlying drivers of forest disturbance frequently co-occur and interact. Timber extraction, for instance, is driven by market demand but is facilitated by corruption and weak governance (*34*); forest fires are often caused by agricultural practices but can be exacerbated by extreme droughts (*23*). Furthermore, there are important and multiscale feedbacks between the drivers of disturbances and their impacts. At the landscape scale, deforestation or degradation-related disturbances cause warming and alter precipitation, potentially increasing drought (*8*, *35*). At the global scale, carbon dioxide emissions from forest disturbance are major contributors to climate change, driving extreme droughts that cause or amplify degradation (*24*, *36*). Anthropogenic disturbances in Amazonian forests are therefore the result of the interplay between a broad suite of drivers that are expressed and interact across a range of spatial scales (Fig. 1). Understanding their impacts is no less complex and requires quantifying the intensity and severity of disturbances and their distribution and interplay over time and space (Box 1).

Spatial extent and severity

Over the past decades, uncertainty in determining the extent of degradation (Box 1) has been minimized by advances in remote-sensing technology. The increased availability of time-series information from the Terra, Aqua (MODIS sensor) and Landsat (TM sensor) satellites has helped demonstrate the widespread occurrence and impact of tropical forest degradation (*12*, *15*, *16*, *37*–*39*). The only existing pan-Amazonian direct estimate using a Landsat time series (*11*) indicates an area of 1,036,800 ± 24,800 km^2 affected by human and natural disturbances between 1995 and 2017 (47,127 ± 1127 km^2 $year^{-1}$), corresponding to 17% of the total forest area in 2017. Disentangling the spatial extent and severity of the multiple drivers of degradation is critical for understanding the impact of disturbances on tropical forests. Each disturbance type is driven by distinct factors, leading to great variation of their spatial extents from year to year. To capture the patterns of multiple disturbances, we compiled published data of the four main drivers of forest degradation, using the most up-to-date, spatially explicit datasets on burned area (*40*), timber extraction (*41*), edge effects (*9*), and drought (*42*). We assessed the period from 2001 to 2018. Data for the four disturbances had spatial resolutions of 0.5, 27, 0.03 and 55 km, respectively. We show that in that period fires alone affected 122,624 km^2, timber extraction 119,700 km^2, edge effects 188,531 km^2, and drought 2,740,647 km^2 (Fig. 2), representing, respectively, 1.8, 1.8, 2.8, and 41.1% of the remaining Amazon forest cover (6,673,908 km^2) (*43*).

Forest fires intensify during drought years (*10*, *23*, *24*, *44*, *45*), leading to acute peaks in burned area: 14,584 and 32,815 km^2 in the dry years of 2005 and 2010, respectively. This is two to four times the mean total forest area burned in all other years in the 2001–2018 period (7701 km^2). Although the extent of Amazonian fires during recent droughts has already been large, much larger megafires are also possible (*46*). Edge creation is strongly and positively correlated with deforestation at the basin level (*9*), although further deforestation could reduce the area of forests exposed to edges in regions with low levels of forest cover.

Despite remaining stable over time in the analyzed dataset, timber extraction extent remains highly uncertain. The product used here (*41*) shows an annual rate of 6623 km^2 $year^{-1}$ affected by timber extraction from 2001 to 2018 in the Brazilian Amazon. The first Brazilian Amazon–wide study estimated a rate of 11,537 km^2 $year^{-1}$ between 1999 and 2002 (*27*), which coincides with a period of high deforestation rates in the region. The other Brazilian Amazon–wide estimate assessing the extent of timber extraction from 1992 to 2014 showed an annual rate of 4479 km^2 $year^{-1}$ (*12*). This is 32% lower than the timber extraction estimate shown in Fig. 2, a difference that may be related to the frequency of the temporal series analyzed by Matricardi *et al.* (*12*), or the difference between the timber extraction product used here—which is based on national census statistics—and a remotely sensed approach. Estimates suggest that ~50% (or even more) of the timber extraction in the Amazon is illegal (*47*), meaning that this does not appear in either the national census or the product used here.

The complexity of quantifying degradation impacts increases with the frequency of overlap among different disturbances. Using the two highest spatial resolution datasets (burned and edge areas), we found that 25% of the total burned forest area was within 120 m of an edge, affecting 17% of the total edge area. Additionally, 6% of the area affected by both edges and fire was also affected by drought. Accounting for the spatial extent of forests hit by fire, timber extraction and edge effects, and the overlaps between them, the degraded area due to these three drivers affected at least 364,748 km^2 (5.5% of all remaining Amazonian forests) from 2001 to 2018. This corresponds to 112% of the total area deforested during this same period (325,975 km^2) (*15*, *48*, *49*), and is within the same magnitude of a previous estimate of degradation in the Brazilian Amazon of 337,427 km^2 in the 1992–2014 period (*12*). Estimating overlap between timber extraction and other drivers is not trivial because the data used for timber extraction provide a percentage cover by area within the 27-km-resolution grid cell, rather than a precise delimitation of the area affected by these events. Here, we assumed a proportional distribution of logged forests within the grid cell to account for the timber extraction overlaps with other drivers, as logged forests can be more flammable than undisturbed forests (*50*), and the extraction of timber is often associated with edges (*51*). Not all the extreme droughts observed in the Amazon

Box 1. Defining Amazonia's degradation regime.

The important factors determining impact can be understood by extending the concept of a fire regime to the disturbances that cause degradation.

Extent: The area of forest affected by disturbances. Severe disturbances that affect canopy cover can be assessed using remote sensing; more subtle changes resulting from droughts can be inferred from anomalies in water deficit (*130*) (see "Spatial extent and severity" section).

Intensity: A measure of the strength of a disturbance, such as logging offtake, fire radiative power, the strength of the water deficit anomaly, or degree of exposure to edges.

Severity: A measure of the impact of the disturbance on ecosystem-level or social conditions. This is a function of disturbance intensity and the sensitivity of the ecosystem or of societal groups that depend on forest resources.

Frequency: The number of disturbance events. The severity of disturbance often increases with the number of disturbance events, and recurrent fires or logging can bring about dramatic changes in ecological condition on decadal time scales.

Co-occurrence: The incidence of different forms of disturbance occurring in the same location (Fig. 2), in part encouraged by the interactions among them (see "Underlying drivers of disturbance" section). Co-occurring disturbances can amplify the severity (e.g., fire effects are more severe near edges). Co-occurrence can also be important at the landscape level, even if disturbances are not precisely superimposed. Their combined effect can contribute to significant losses of biodiversity (*7*) and of ecosystem services that are valuable to human populations (see "Social and economic dimensions" section).

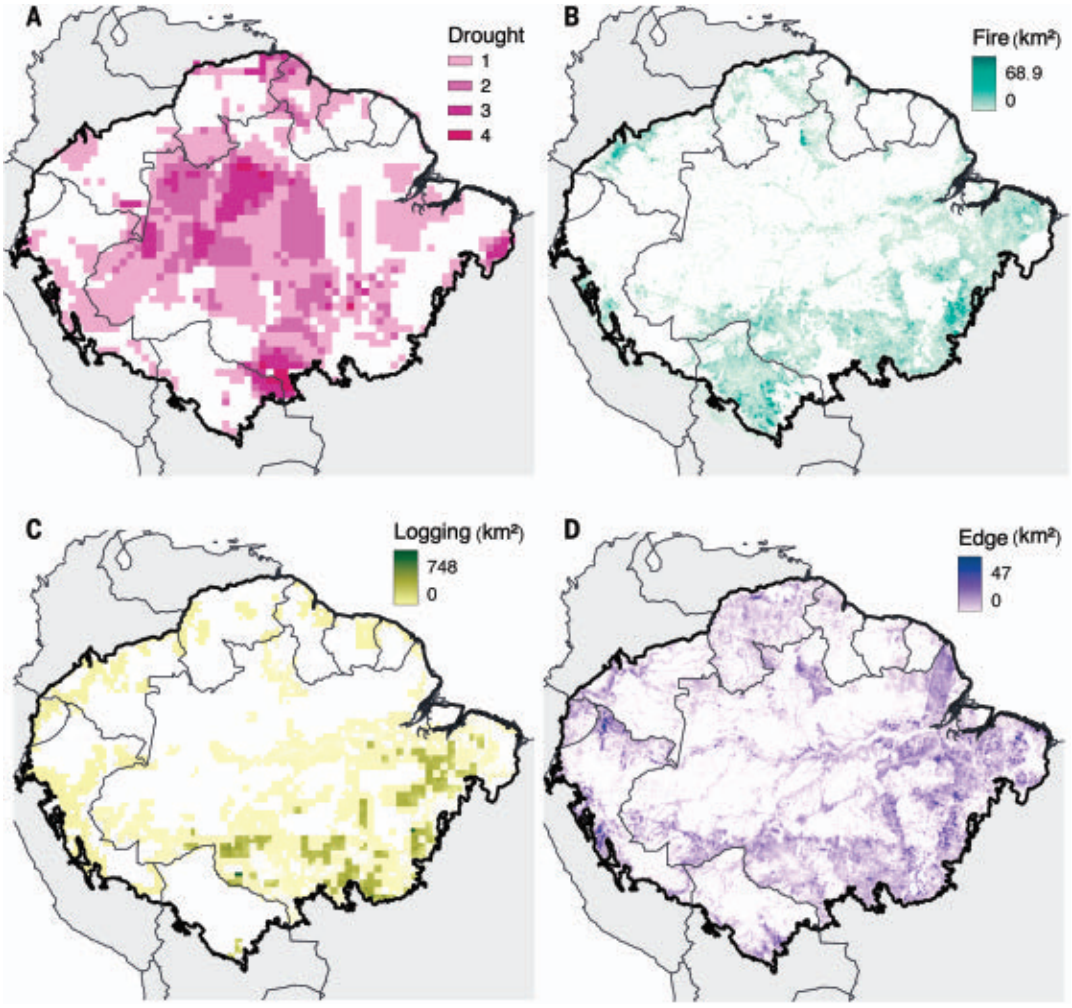


Fig. 2. Current (2001–2018) spatial distribution of the four main drivers of forest degradation in the Amazon forest, excluding deforestation and savanna areas. (**A**) Extreme drought occurrence, (**B**) burned area, (**C**) timber extraction, (**D**) area within a forest edge. The datasets employed (*23*, *40*–*42*, *129*), processing steps, and numerical estimates are shown in the supplementary materials.

in the analyzed 2001–2018 period have been unequivocally attributed to human-induced climate change (*22*). Nevertheless, when considering all the four main drivers, and all possible overlaps between them, the estimate of total degraded area increases to 2,542,593 km^2, or 38% of the remaining Amazonian forests. This total degraded area includes 628,909 km^2 of forest where two or more of the four disturbances overlap (table S1).

This assessment indicates a broad range of estimates, varying from 5.5% (considering only fire, timber extraction, and edge effects) to 38% (considering all four disturbances). Such a large range of estimates of the extent of degradation is mainly determined by the types of disturbances considered (with much larger area estimates if less-severe disturbances, such as droughts, are included; Fig. 2), the spatial and temporal ranges of the studies, and their distinct methods (*16*, *37*–*39*, *52*). All recent studies, however, consistently agree that the extent of degraded forest is growing, and the total area is either equal to or greater than the Amazon's deforested area (*10*–*12*).

Ecological impacts

Changes in carbon stocks and basin-wide emissions

Disturbance type and intensity are strong predictors of the magnitude of change in aboveground carbon stocks (i.e., severity; Box 1). Carbon losses are often greatest in burned forests, compared to the other disturbances (*53*). Sixty-nine percent of the burned area shown in Fig. 2 has been affected by a single understory forest fire, reducing aboveground carbon stocks by 13 to 50% (*17*, *54*, *55*) (Fig. 3). Tree mortality following understory fires varies spatially: The highest levels of tree mortality and the greatest biomass losses have been recorded in the Brazilian state of Pará (*29*, *56*). Smaller effects have been recorded in drier Amazonian regions (*44*) where trees are protected by thicker bark (*57*) and in less-seasonal regions where fire intensity may be limited by high fuel moisture content (*54*). Carbon losses in logged forests are also highly variable and range from 4 to 35% (Fig. 3), depending on extraction intensity and the management of collateral damage (*28*). The severity of edge effects varies in relation to the distance to the forest edge, with severity decreasing from the edge to the interior, and over time, with most losses occurring within 5 years of edge formation. Even when these factors are controlled, the impacts vary substantially: Carbon losses within 120 m of an edge range from 23 to 35% in the first 4 years after the edge formation (*6*, *26*), with the severity potentially related to exposure to fire (*9*) (Fig. 3). Edge effects may also vary over much larger spatial scales and could be less severe in forests on the richer soils of western Amazonia (*58*). Finally, extreme droughts bring about short-term carbon losses of 1 to 8% (*23*, *59*) (Fig. 3).

Time since disturbance is an important determinant of aboveground carbon stocks. When forests are burned, the recovery of carbon stocks from tree recruitment and growth is offset by

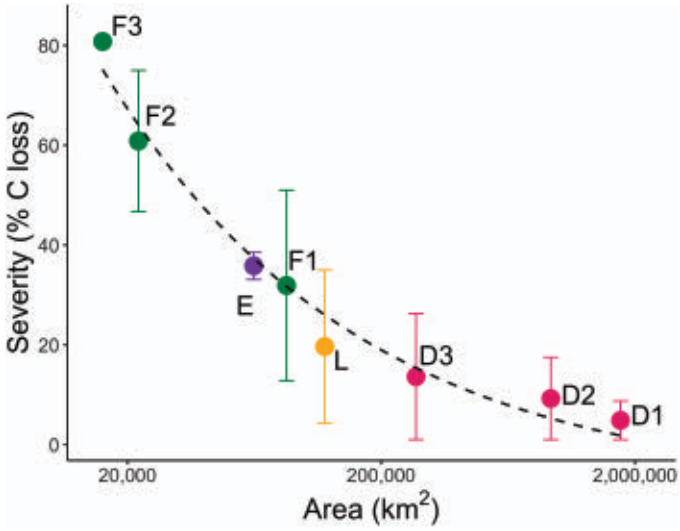


Fig. 3. The relationship between the area affected between 2001 and 2018 and disturbance severity (carbon loss). D, extreme drought; E, edge effects; F, forest fire; L, timber extraction (logging). Numbers denote single events (1) or repeated fires or droughts (2 or 3). Area is shown on a $\log_{10}$ scale. See supplementary materials for analysis methods.

high rates of ongoing tree mortality (*23*, *56*, *60*), such that burned forest can be a net source of carbon emissions for up to 7 years after the fire and hold ~25% less carbon after 30 years (*45*, *55*, *60*). Biomass recovery times after logging are almost directly proportional to the volume of timber extracted, such that extraction of 10, 25, or 50% of prelogging aboveground carbon stocks would require 12, 43, or 75 years to recover (*28*). These rates also vary across the Amazon, depending on soil fertility and climate (*50*). Carbon losses from edge effects are most pronounced after the first 4 years (*9*, *26*). As 66% of current edges are older than this (*6*), most will have incurred these losses. Longer-term assessments of drought impacts show mixed results, with plot-based studies reporting both rapid recovery (*61*) and sustained effects lasting at least 3 years (*23*).

Repeated disturbances are often associated with the greatest losses of aboveground carbon. Recurrent fires can lead to losses of over 80% of aboveground carbon (*17*) (Fig. 3), which is important as almost one-third of the burned area has been burned either twice (18%) or three or more times (13%) (Fig. 2). Similarly, the impacts of timber extraction are far greater in forests that have suffered multiple extraction events (*53*), and edge effects are greater when forests are exposed to multiple edges (*62*). The cumulative impact of multiple droughts on aboveground carbon is not known but could be important given that over one-third of the drought-affected area was affected by two (26%) or more (10%) events in an 18-year period (Fig. 2). Co-occurring disturbances can also amplify effects, with windstorms resulting in much higher biomass losses in thrice-burned forest (31%) than in unburned forests (15%) (*30*).

Our overview of the extent (Fig. 2), severity (Fig. 3), and longevity of these four disturbances demonstrates that they are likely to be a substantial source of long-term carbon emissions from Amazonian forests. However, at present there is insufficient information on disturbance recurrence and recovery to make a reliable estimate of their combined influence on the Amazon's carbon balance. Studies that have attempted this using remote sensing, or mixing field assessments with estimates of extent, estimate annual emissions of between 0.05 and 0.2 Pg C $year^{-1}$ for a different combination of disturbances (*10*–*16*), which are comparable to deforestation emission estimates of 0.06 to 0.21 Pg C $year^{-1}$ (*49*, *63*). Yet comparisons remain confounded by the different spatial and temporal scales of assessments and the different types of disturbance that are being assessed—studies inferring degradation from canopy openness are likely to miss some of the degradation resulting from edges or low-intensity logging, while airborne air sampling is unable to accurately separate emissions from deforestation and degradation (*8*).

Other climate processes

Beyond carbon, forest disturbances influence a range of atmospheric processes. Within the forests themselves, tree mortality from forest fires, timber extraction, and edge effects increase temperatures and lower the humidity of the understory (*26*, *29*, *35*, *50*). Reductions in forest biomass and changes in species composition also affect water cycling (*64*). Forest edges generate 5% less evapotranspiration (ET) than forest interiors (*65*), and degraded forests provide between 2 and 34% less ET than intact forests in normal dry seasons, with stronger reductions in southern drier sites (*35*). However, the magnitude of change for ET seems to be far less than that for carbon stocks (*35*, *66*), with recovery occurring within 7 years of repeated forest fires (*66*). Amazonian fires also reduce air quality many thousands of kilometers from the source (*67*), while soot deposits are accelerating glacier melt in the Andes (*68*).

Biodiversity and ecosystem functioning

Fires, timber extraction, and edge effects reduce the number of forest species (*7*, *69*) and species with the highest conservation values (*7*). In landscapes with ~80% forest cover in the eastern Amazon, the combined influence of forest disturbances in remaining forest results in about as much biodiversity loss as the loss of habitat in the deforested areas (*7*). In fragmented landscapes, patch area is an important determinant of species persistence; conserving the full suite of forest birds requires maintaining large patches (e.g., >10,000 ha) of good-condition forest (*70*). The impacts of forest disturbance extend to aquatic biota, and even reduced-impact logging affects the composition and functional traits of stream fishes (*71*). Disturbance also disrupts multitrophic processes such as pollination, decomposition, seed dispersal, and herbivory (*72*) and drives sublethal changes in the morphology or physiology of birds (*73*) and dung beetles (*69*).

The postdisturbance recovery of forest fauna can be slow, and understory forest-specialist birds do not recover their original abundances even 10 years after a single fire event (*74*). Recovery can be further impeded where forests have been affected by previous disturbances (*72*) or where succession is dominated by lianas, palms, bamboos, or invasive grasses (*45*, *75*). Finally, fauna can support postdisturbance forest recovery, with birds, terrestrial ungulates, and primates all helping disperse seeds (*76*). Some of these taxa are resilient to low-intensity disturbance (*77*) and contribute to forest regeneration where hunting is controlled and there is connectivity with undisturbed forests (*78*).

Social and economic dimensions

Whether initiated by chainsaws, fire, or drought, people drive forest degradation (*21*). Its prevalence and persistence (Fig. 2) are largely explained by the (short-term) economic benefits driving the four disturbances (Fig. 4A). A broad range of human actors generate these disturbances, from local forest communities that use fire for subsistence agriculture, regional commercial businesses extracting timber, to distant city-dwellers consuming commodities originating in forest landscapes, the fossil fuel consumption of the international community, and investment banks contributing to geopolitical and market forces (*79*). These actors are influenced by the outcomes of disturbances (e.g., smoke from fires), and the resulting degraded

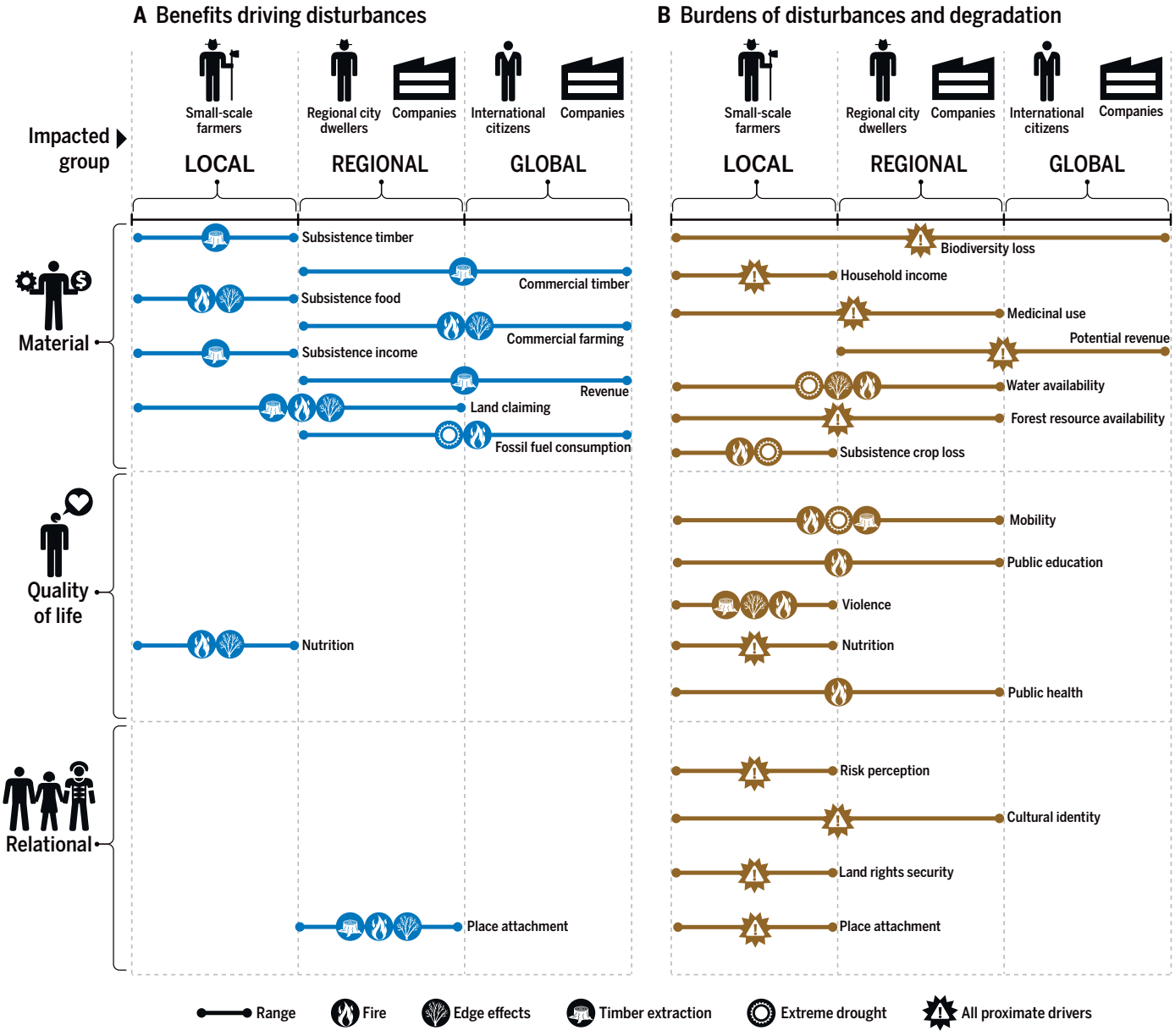


Fig. 4. Socioeconomic benefits and burdens of Amazon forest degradation and its drivers are unevenly distributed. (**A**) The underlying drivers of disturbance generate mainly material benefits, while (**B**) the resulting forest degradation generates burdens that are unevenly distributed among stakeholders and across scales. Impacts are displayed as benefits (blue) and burdens (brown) to people, with drivers and outcomes grouped according to material, quality of life, and relational dimensions. The main disturbance type(s) associated with each outcome is indicated by the icon and the range of impacts by the horizontal extent. This list is not exhaustive.

forest states, in distinct [i.e., material, subjective (quality of life), and relational impacts] and unevenly distributed ways across multiple spatial scales (Fig. 4). Crucially, the flow of benefits (which are related to proximate drivers of forest degradation) and burdens (which are related to degradation outcomes) is misaligned. Benefits often accrue to external stakeholders, whereas burdens are concentrated locally, creating socioecological injustices (Fig. 4B). To achieve more just and sustainable outcomes, the benefit seeking that ultimately drives degradation needs to be balanced against the multitude of burdens that arise from it.

Material benefits of degradation

Many of the disturbances driving degradation deliver material benefits (i.e., money and goods) to privileged elites living outside of forest landscapes (*80*, *81*). For example, a small fraction of (large-scale) landholders account for most forest loss (*82*), contributing to degradation via edge creation through deforestation, escaped pasture renewal fires (*32*), and reduced regional rainfall (*8*) and ultimately contributing to global climate change itself via carbon emissions (*83*). Forest loss is strongly associated with commodity production, with material benefits accruing to wealthy regional and international actors (*84*). For instance, even omitting the sizable clandestine market (*34*, *47*), timber extraction in the Brazilian Amazon generated US$459.5 million in 2018 (*85*), which was not well distributed (*4*).

At local scales, small-holder farmers contribute to forest degradation, either directly through small-scale timber extraction or hunting, or indirectly via agricultural fires that

CREDIT: ALEX ARGOZINO/STUDIO ARGOZINO

may escape into forests. Notably, benefits of local drivers are retained locally (e.g., supporting household incomes) or regionally (e.g., food security) (*86*) (Fig. 4A). Other benefits accruing locally include income from hired labor in logging camps (*87*). However, these economic benefits tend to be short term (*4*), poorly negotiated, and disproportionately small and do not compensate for the local damages that forest disturbances inflict (*88*, *89*).

Multidimensional burdens exacerbate the vulnerabilities of marginalized groups

Forest disturbances driving degradation create burdens to multidimensional human well-being that are predominantly concentrated on local communities (Fig. 4B). The most severe material impacts are borne by small-scale farmers, Indigenous people, and traditional communities who rely on a diverse set of forest resources to underpin resilient livelihoods and cultural practices (*90*). Timber extraction reduces the availability of species that contribute to nutritional diversity or provide oils or medicines (*89*, *90*). Reductions in diversity of host species undermine the "dilution effect" (where low-quality host abundance buffers parasite dispersal), increasing vector-borne diseases (e.g., Chagas disease) (*91*). The dense understory of fire-affected forests makes hunting harder (mobility in Fig. 4B) and reduces availability of preferred game species (*92*). Forest disturbances related to degradation can alter fish abundance in streams, rivers, and floodplains, with implications for the nutritional diversity and food security of local communities (*71*). Some of the material changes extend beyond local communities, as reductions in forest resources can affect peri-urban households that maintain strong links with forests (*88*), compounding existing vulnerabilities associated with structural marginalization (*93*).

Although less well understood than material impacts, burdens related to forest disturbance also affect the relational and subjective dimensions of people's lives, which make important contributions to human well-being (*94*). Further, some of the disturbances that cause degradation (e.g., burning, presence of logging operations) themselves reduce the quality of life of local peoples—for example, by increasing the exposure of forest peoples to infection (e.g., COVID-19) (*95*). Public-health burdens accrue from the smoke associated with fires and include premature deaths as well as school closures that potentially reduce the learning lifetime of local children (public education in Fig. 4B) (*67*, *96*, *97*). Incidence of violence rises when land conflicts associated with forest degradation occur (*98*), and within temporary settlements created for logging operations (*99*).

The loss of forest resources following disturbances can negatively influence relational dimensions of people's lives, including sociocultural reproduction, cohesion, and cultural practices. For example, forest degradation can erode communal sites, impair place attachments, and affect interactions with the forest and ways of knowing and of using its resources. Degradation can also heighten perceptions of vulnerability and risk owing to place dislocation, transformation, and threat of potential resettlement (*100*).

Diffuse and indirect burdens accrue to external actors

Amazon forest disturbances that drive degradation also burden regional and international actors, though often in more indirect and diffuse ways. For instance, people living large distances from forests may be affected by disturbance-induced changes in the carbon and water cycle (*35*). These impacts extend to regions surrounding the Amazon, with implications for material gains and revenue within the agriculture sector (*83*). Fires can influence the sustainability of water availability in distant (e.g., Andean) cities, with the deposition of black carbon accelerating glacial melt (*68*). Fire also causes material damage to the timber sector, affecting commercially unexplored forests (*17*, *101*), and other losses to potential revenue and to the region's economy (e.g., through airport closures) (*96*).

Forest degradation precludes discovery of new pharmaceutical, nutritional, and bio-based products and can precipitate the emergence of pandemics with global consequences for health, economies, and well-being (*102*, *103*). Further, the relationship between ecosystem degradation and regional public health has the potential to be important (*104*). Estimates suggest that the loss of ecosystem services as a result of extreme climate change in the Amazon may induce regional economy losses of US $7.7 trillion in a period of 30 years (*18*), and this excludes the substantial intangible relational and quality-of-life impacts. Better understanding of the multifaceted suite of burdens extending from degradation across scales could help to inform appropriate policy responses and galvanize support in society for a shift toward more sustainable use of the forest.

Projecting Amazon forest degradation

Most studies assessing future scenarios for the Amazon focus on deforestation and its relationship with prospective road development, agricultural expansion, and conservation policies (*3*, *105*–*108*). Only five studies have projected future Amazon forest degradation in a spatially explicit way, either covering the entire Amazon biome (*109*) or focusing on the Colombian (*110*), Brazilian (*111*, *112*), or southern Brazilian Amazon (*44*). Modeling approaches include mechanistic (*111*), statistical (*110*, *112*) and hybrid (*44*, *109*) methods. Studies assessing the proximate causes (Fig. 2) focused on fire occurrence driven by deforestation and climate change (*44*, *109*), fire intensity driven by climate change (*111*), edge effects due to forest fragmentation (*110*), or mixed causes (*112*). Two further studies modeled degradation in a nonspatially explicit way using fixed degradation-to-deforestation ratios (*113*) or statistical relationships of carbon loss caused by logging and fire (*114*). Despite the variation in methods and study areas, these modeling studies reinforce many of the findings emerging from empirical studies, including that (i) feedbacks between different drivers are key for Amazon forest degradation (*109*); (ii) degradation can occur independently from deforestation [e.g., control of deforestation can reduce fire activity, but only under weak to moderate climate change scenarios (*109*)]; (iii) climate change can boost fire intensity and ignition sources, promoting fire-driven degradation (*111*); (iv) roads promote degradation as well as deforestation (*51*); and (v) carbon dioxide emissions from degradation can overwhelm those from deforestation (*44*, *110*), and the carbon uptake from regeneration (*113*).

Combining previously published projections of the individual main disturbances that cause degradation (*43*), we project potential future patterns of degradation of the Amazon forest and their effects on carbon stocks under two alternative deforestation scenarios: "governance" (GOV) and "business-as-usual" (BAU). These projections show (Fig. 5) that halting deforestation, as pledged by Amazonian nations in the Glasgow declaration and in their nationally determined contributions to the Paris Agreement, does not necessarily curb degradation across the Amazon. Projected 2050 annual carbon emissions are 0.06 Pg C $year^{-1}$ in the GOV scenario and 0.42 Pg C $year^{-1}$ in the BAU scenario. This upper limit is considerably higher than the upper limit of 0.2 Pg C $year^{-1}$ observed in the 2001–2018 period, owing to a stronger contribution of more frequent droughts in the future, but still lower than another projection restricted to the Brazilian Amazon (*112*). Indeed, the degradation-to-deforestation ratio for carbon emissions remains high both in a scenario where illegal deforestation is stopped after 2030 (GOV, 1.04) and in the BAU scenario, which extrapolates the land-use dynamics of the early 2000s (BAU, 0.74). To some extent, these findings are to be expected, given that halting deforestation leaves a larger forest area that is subject to fires, logging, or droughts (*115*). However, our assessment also indicates the importance of designing and implementing intervention strategies that address degradation and deforestation as distinct processes (*116*, *117*).

Emissions in the GOV scenario are dominated by fire (59%), followed by droughts (38%) and logging (3%). However, in the BAU scenario, under Representative Concentration Pathway

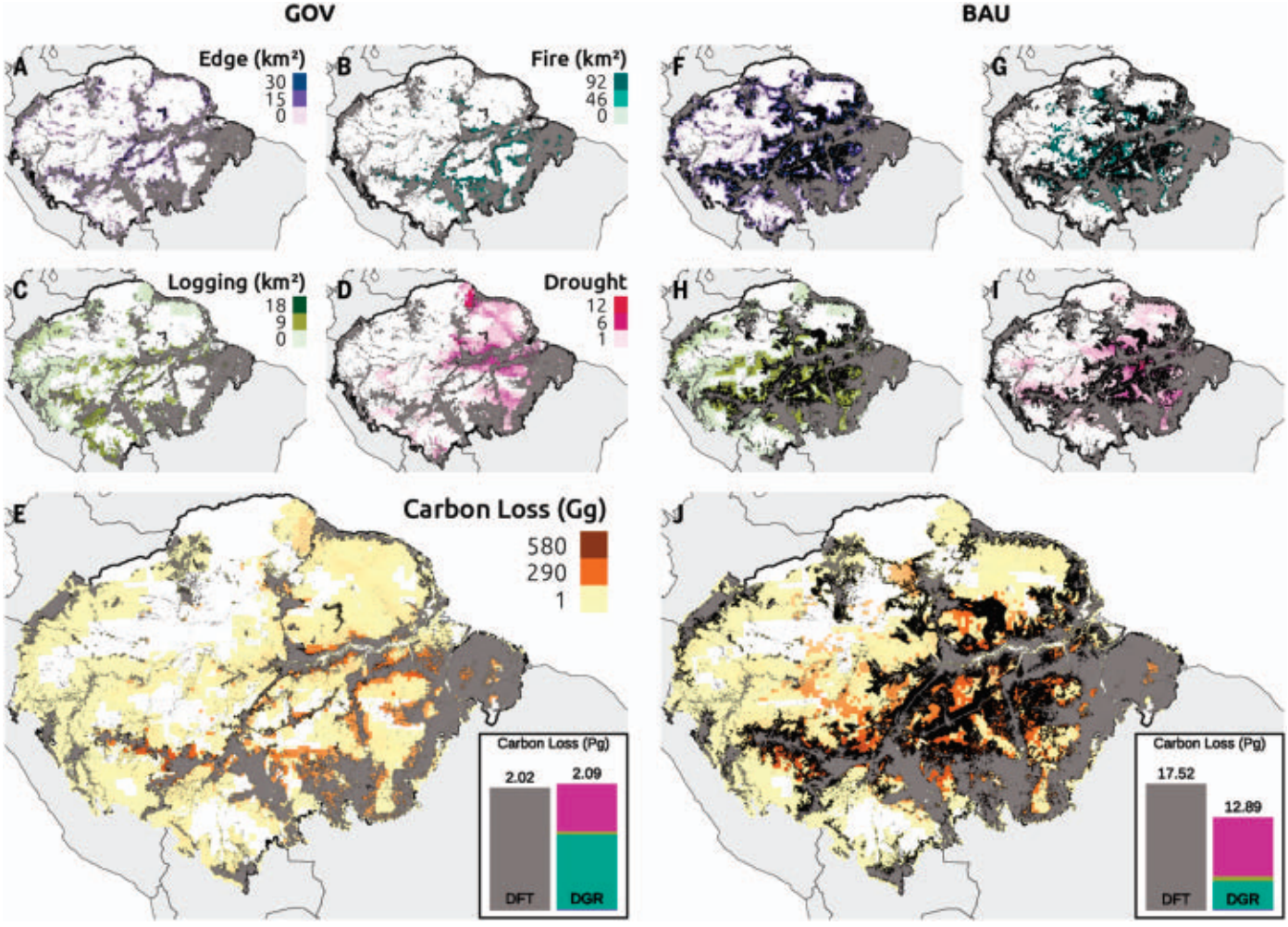


Fig. 5. First-order 2050 projections of Amazon forest degradation through its main drivers. Projections of 2019–2050 changes in the main proximate drivers of Amazon forest degradation. (**A** and **F**) Edge effects (*108*); (**B** and **G**) fire occurrence (*111*); (**C** and **H**) timber extraction (*41*); (**D** and **I**) extreme drought (in number of occurrences in 2019–2050) (*33*); and resulting combined carbon losses (**E** and **J**) under climate and deforestation governance (GOV) and business-as-usual (BAU) scenarios (*43*, *108*). Inset charts in (E) and (J) show resulting carbon emissions in the 2019–2050 period resulting from deforestation (DFT) and degradation (DGR) (notice the different scales). The share of C emissions per driver is shown in the DGR bar and follows map colors. Black map areas denote deforestation in the 2019–2050 period, whereas gray areas depict deforestation prior to 2019. See supplementary materials for methods and numerical results.

(RCP) 8.5 climate, droughts become the dominant cause of carbon emissions associated with degradation (63%), followed by fire (30%) and logging (5%). The high relative contribution of drought demonstrates that the mitigation of Amazon forest degradation also depends on concerted international (i.e., extra-Amazonian) efforts to abate global climate change. These findings are aligned with observational data regarding the hierarchy of each disturbance in terms of carbon loss and affected area (Fig. 3 and table S2).

Although these projections demonstrate the potential importance of future degradation, they have probably been underestimated as they do not include feedbacks and interactions between disturbances (see "Underlying drivers of disturbance" section). For example, degradation from timber extraction, extreme droughts, and edge effects alter the forest microclimate, making future fires more likely (*29*). The feedbacks between Amazon forest degradation and regional climate change are particularly relevant for determining the likelihood of an Amazon tipping point (*118*, *119*).

Degradation and the future of the forest

Although our understanding of degradation has improved markedly, important uncertainties remain regarding the quantification of the area affected by the different disturbances, their longer-term impacts, and how disturbance severity is modified by co-occurrence, repeated events, or changes in management practices (e.g., toward integrated fire management, or sustainable logging protocols). Our understanding of the drivers of disturbance would be improved by more in-depth analyses of underlying causes and better identification of the actors and funding chains, as has been extensively investigated for deforestation (*3*–*5*, *82*). Further, research is essential into what forms of governance, co-responsibility, and valuation can best, and most realistically, balance the environmental, social, and economic imperatives associated with forest resource management (*120*, *121*).

From the policy perspective, the distinct nature of proximate drivers, the range of stakeholders that benefit from them, and the challenges in monitoring disturbances all make curbing forest degradation considerably more complex than reducing deforestation. The Reduction of Emissions from Deforestation and Degradation (REDD+) framework is the only existing international policy mechanism that aims to address tropical forest degradation (*6*). Nevertheless, only a small minority

of REDD+ projects are targeted at preventing degradation (*117*), and the identification of key actors and drivers in REDD+ projects is confusing, even when those address the well-known process of deforestation (*122*). Moreover, although leakage effects (displacement of deforestation from a REDD+ covered area to another area not covered by that program) are a major concern for deforestation-based projects (*123*), they remain unquantified for displaceable disturbances such as timber extraction.

Although our intent here is not to be policy prescriptive, this Review has nonetheless revealed some key priorities for policy-makers and practitioners. Preventing further deforestation remains a key objective for stabilizing the climate system, preserving biodiversity, and ensuring sustainable development; deforestation is itself a major driver of greenhouse gas emissions and biodiversity loss and a driver of several forms of degradation (Fig. 1). The integrity of the basin also depends on maintaining sufficient forest cover (*119*). Preventing additional degradation will also benefit from the conditions required to curb deforestation, such as the strengthening of land tenure, environment-oriented credit concession, and the provision of sustainable income and livelihood alternatives that can attenuate social inequalities (*124*).

But it is also clear that actions taken to prevent deforestation are not enough and must be supported by other interventions, such as preventing illegal logging (*34*), implementing large-scale investments and capacity building for a shift to fire-free cattle ranching, and supporting smallholders to reduce, eliminate, or better control the use of fires in agriculture. Initiatives to curb degradation (and stimulate restoration) arising from the private sector should be encouraged by public policies, learning from previous initiatives such as the efforts to avoid deforestation in the Amazonian soybean production sector (*125*, *126*). All these actions will benefit from improvements in the monitoring of tropical forest degradation. As spaceborne light-detecting and ranging (LiDAR) technology becomes increasingly cost-effective (*127*), the combination of its ability to detail canopy structure with optical imagery is a promising avenue for operationalizing the monitoring of disturbances linked to degradation (*128*). Other innovative ground-based monitoring initiatives such as the "smart forests" concept could be useful in contexts where disturbances such as timber extraction are key threats (an example is given by the Rainforest Connection initiative, https://rfcx.org/). Finally, efforts to reduce degradation will all be supported by rapid and multidisciplinary advances in our socioenvironmental understanding of tropical forest degradation that can provide a robust platform on which to co-construct appropriate policies and programs to curb it (*6*).

REFERENCES AND NOTES

1. Y. Malhi, T. A. Gardner, G. R. Goldsmith, M. R. Silman, P. Zelazowski, Tropical Forests in the Anthropocene. *Annu. Rev. Environ. Resour.* **39**, 125–159 (2014). doi: 10.1146/annurev-environ-030713-155141
2. X.-P. Song *et al*., Global land change from 1982 to 2016. *Nature* **560**, 639–643 (2018). doi: 10.1038/s41586-018-0411-9; pmid: 30089903
3. D. Nepstad *et al*., Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. *Science* **344**, 1118–1123 (2014). doi: 10.1126/science.1248525; pmid: 24904156
4. A. S. L. Rodrigues *et al*., Boom-and-bust development patterns across the Amazon deforestation frontier. *Science* **324**, 1435–1437 (2009). doi: 10.1126/science.1174002; pmid: 19520958
5. H. J. Geist, E. F. Lambin, Proximate Causes and Underlying Driving Forces of Tropical Deforestation: Tropical forests are disappearing as the result of many pressures, both local and regional, acting in various combinations in different geographical locations. *Bioscience* **52**, 143–150 (2002). doi: 10.1641/0006-3568(2002)052[0143:PCAUDF]2.0.CO;2
6. C. H. L. Silva Junior *et al*., Amazonian forest degradation must be incorporated into the COP26 agenda. *Nat. Geosci.* **14**, 634–635 (2021). doi: 10.1038/s41561-021-00823-z
7. J. Barlow *et al*., Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature* **535**, 144–147 (2016). doi: 10.1038/nature18326; pmid: 27362236
8. L. V. Gatti *et al*., Amazonia as a carbon source linked to deforestation and climate change. *Nature* **595**, 388–393 (2021). doi: 10.1038/s41586-021-03629-6; pmid: 34262208
9. C. H. L. Silva Junior *et al*., Persistent collapse of biomass in Amazonian forest edges following deforestation leads to unaccounted carbon losses. *Sci. Adv.* **6**, eaaz8360 (2020). doi: 10.1126/sciadv.aaz8360; pmid: 32998890
10. L. E. O. C. Aragão *et al*., 21st Century drought-related fires counteract the decline of Amazon deforestation carbon emissions. *Nat. Commun.* **9**, 536 (2018). doi: 10.1038/s41467-017-02771-y; pmid: 29440640
11. E. L. Bullock, C. E. Woodcock, C. Souza Jr., P. Olofsson, Satellite-based estimates reveal widespread forest degradation in the Amazon. *Glob. Chang. Biol.* **26**, 2956–2969 (2020). doi: 10.1111/gcb.15029; pmid: 32022338
12. E. A. T. Matricardi *et al*., Long-term forest degradation surpasses deforestation in the Brazilian Amazon. *Science* **369**, 1378–1382 (2020). doi: 10.1126/science.abb3021; pmid: 32913104
13. S. L. Maxwell *et al*., Degradation and forgone removals increase the carbon impact of intact forest loss by 626. *Sci. Adv.* **5**, eaax2546 (2019). doi: 10.1126/sciadv.aax2546; pmid: 31692892
14. Y. Qin *et al*., Carbon loss from forest degradation exceeds that from deforestation in the Brazilian Amazon. *Nat. Clim. Chang.* **11**, 442–448 (2021). doi: 10.1038/s41558-021-01026-5
15. A. Tyukavina *et al*., Types and rates of forest disturbance in Brazilian Legal Amazon, 2000-2013. *Sci. Adv.* **3**, e1601047 (2017). doi: 10.1126/sciadv.1601047; pmid: 28439536
16. W. S. Walker *et al*., The role of forest conversion, degradation, and disturbance in the carbon dynamics of Amazon indigenous territories and protected areas. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 3015–3025 (2020). doi: 10.1073/pnas.1913321117; pmid: 31988116
17. J. Barlow *et al*., The critical importance of considering fire in REDD+ programs. *Biol. Conserv.* **154**, 1–8 (2012). doi: 10.1016/j.biocon.2012.03.034
18. D. M. Lapola *et al*., Limiting the high impacts of Amazon forest dieback with no-regrets science and policy action. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 11671–11679. doi: 10.1073/pnas.1721770115
19. J. A. Parrotta *et al*., "Conclusions" in *Understanding relationships between biodiversity, carbon, forests and people: The key to achieving REDD+ objectives*, J. A. Parrotta, C. Wildburger, S. Mansourian, Eds. (IUFRO, Vienna, 2012), pp. 139–146.
20. L. Olsson *et al*., "Land Degradation" in *Climate Change and Land: An IPCCSpecial Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*, P. R. Shukla *et al*., Eds. (2019), pp. 345–436; https://doi.org/10.1017/9781009157988.006.
21. J. Ghazoul, R. Chazdon, Degradation and Recovery in Changing Forest Landscapes: A Multiscale Conceptual Framework. *Annu. Rev. Environ. Resour.* **42**, 161–188 (2017). doi: 10.1146/annurev-environ-102016-060736
22. G. G. Ribeiro Neto *et al*., Attributing the 2015/2016 Amazon basin drought to anthropogenic influence. *Clim. Resil. Sustain.* **1**, e25 (2022). doi: 10.1002/cli2.25
23. E. Berenguer *et al*., Tracking the impacts of El Niño drought and fire in human-modified Amazonian forests. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2019377118 (2021). doi: 10.1073/pnas.2019377118; pmid: 34282005
24. S. L. Lewis, P. M. Brando, O. L. Phillips, G. M. F. van der Heijden, D. Nepstad, The 2010 Amazon drought. *Science* **331**, 554 (2011). doi: 10.1126/science.1200807; pmid: 21292971
25. K. Withey *et al*., Quantifying immediate carbon emissions from El Niño-mediated wildfires in humid tropical forests. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **373**, 20170312 (2018). doi: 10.1098/rstb.2017.0312; pmid: 30297469
26. W. F. Laurance *et al*., An Amazonian rainforest and its fragments as a laboratory of global change. *Biol. Rev. Camb. Philos. Soc.* **93**, 223–247 (2018). doi: 10.1111/brv.12343; pmid: 28560765
27. G. P. Asner *et al*., Selective logging in the Brazilian Amazon. *Science* **310**, 480–482 (2005). doi: 10.1126/science.1118051; pmid: 16239474
28. E. Rutishauser *et al*., Rapid tree carbon stock recovery in managed Amazonian forests. *Curr. Biol.* **25**, R787–R788 (2015). doi: 10.1016/j.cub.2015.07.034; pmid: 26394096
29. M. A. Cochrane *et al*., Positive feedbacks in the fire dynamic of closed canopy tropical forests. *Science* **284**, 1832–1835 (1999). doi: 10.1126/science.284.5421.1832; pmid: 10364555
30. D. V. Silverio *et al*., Fire, fragmentation, and windstorms: A recipe for tropical forest degradation. *J. Ecol.* **107**, 656–667 (2019). doi: 10.1111/1365-2745.13076
31. FAO, "Terms and definitions" in *Global Forest Resources Assessment*, FAO, Ed. (FAO, Rome, 2020), pp. 1–26.
32. J. Barlow, E. Berenguer, R. Carmenta, F. França, Clarifying Amazonia's burning crisis. *Glob. Chang. Biol.* **26**, 319–321 (2020). doi: 10.1111/gcb.14872; pmid: 31729092
33. P. B. Duffy, P. Brando, G. P. Asner, C. B. Field, Projections of future meteorological drought and wet periods in the Amazon. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13172–13177 (2015). doi: 10.1073/pnas.1421010112; pmid: 26460046
34. P. H. S. Brancalion *et al*., Fake legal logging in the Brazilian Amazon. *Sci. Adv.* **4**, eaat1192 (2018). doi: 10.1126/sciadv.aat1192; pmid: 30116781
35. M. Longo *et al*., Impacts of degradation on water, energy, and carbon cycling of the Amazon tropical forests. *J. Geophys. Res. Biogeosci.* **125**, JG005677 (2020). doi: 10.1029/2020JG005677; pmid: 32999796
36. R. Ranasinghe *et al*., "Climate Change Information for Regional Impact and for Risk Assessment" in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, V. Masson-Delmotte *et al*., Eds. (Cambridge Univ. Press, 2021).
37. A. Baccini, W. Walker, L. Carvalho, M. Farina, R. A. Houghton, Response to Comment on "Tropical forests are a net carbon source based on aboveground measurements of gain and loss". *Science* **363**, eaat1205 (2019). doi: 10.1126/science.aat1205; pmid: 30630897
38. A. Baccini *et al*., Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science* **358**, 230–234 (2017). doi: 10.1126/science.aam5962; pmid: 28971966
39. M. C. Hansen, P. Potapov, A. Tyukavina, Comment on "Tropical forests are a net carbon source based on aboveground measurements of gain and loss". *Science* **363**, eaar3629 (2019). doi: 10.1126/science.aar3629; pmid: 30630898
40. L. Giglio, L. Boschetti, D. P. Roy, M. L. Humber, C. O. Justice, The Collection 6 MODIS burned area mapping algorithm and product. *Remote Sens. Environ.* **217**, 72–85 (2018). doi: 10.1016/j.rse.2018.08.005; pmid: 30220740
41. G. C. Hurtt *et al*., Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. *Geosci. Model Dev.* **13**, 5425–5464 (2020). doi: 10.5194/gmd-13-5425-2020
42. I. Harris, T. J. Osborn, P. Jones, D. Lister, Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data* **7**, 109 (2020). doi: 10.1038/s41597-020-0453-3; pmid: 32246091
43. Materials and methods are available as supplementary materials.
44. P. M. Brando *et al*., The gathering firestorm in southern Amazonia. *Sci. Adv.* **6**, eaay1632 (2020). doi: 10.1126/sciadv.aay1632; pmid: 31950083
45. S. S. da Silva *et al*., Impact of fires on an open bamboo forest in years of extreme drought in southwestern Amazonia.

Reg. Environ. Change **20**, 127 (2020). doi: 10.1007/s10113-020-01707-5
46. S. Pueyo *et al*., Testing for criticality in ecosystem dynamics: The case of Amazonian rainforest and savanna fire. *Ecol. Lett.* **13**, 793–802 (2010). doi: 10.1111/j.1461-0248.2010.01497.x; pmid: 20528900
47. E. A. R. Carvalho Jr., E. N. Mendonça, A. Martins, T. Haugaasen, Effects of illegal logging on Amazonian medium and large-sized terrestrial vertebrates. *For. Ecol. Manage.* **466**, 118105 (2020). doi: 10.1016/j.foreco.2020.118105
48. PRODES-INPE, Satellite monitoring of the Amazon forest (2022), (available at http://www.obt.inpe.br/prodes).
49. A. Tyukavina *et al*., Aboveground carbon loss in natural and managed tropical forests from 2000 to 2012. *Environ. Res. Lett.* **10**, 074002 (2015). doi: 10.1088/1748-9326/10/7/074002
50. C. Piponiot *et al*., Carbon recovery dynamics following disturbance by selective logging in Amazonian forests. *eLife* **5**, e21394 (2016). doi: 10.7554/eLife.21394; pmid: 27993185
51. R. Walker *et al*., Modeling spatial decisions with graph theory: Logging roads and forest fragmentation in the Brazilian Amazon. *Ecol. Appl.* **23**, 239–254 (2013). doi: 10.1890/11-1800.1; pmid: 23495649
52. E. L. Bullock, C. E. Woodcock, Carbon loss and removal due to forest disturbance and regeneration in the Amazon. *Sci. Total Environ.* **764**, 142839 (2021). doi: 10.1016/j.scitotenv.2020.142839; pmid: 33131878
53. E. Berenguer *et al*., A large-scale field assessment of carbon stocks in human-modified tropical forests. *Glob. Chang. Biol.* **20**, 3713–3726 (2014). doi: 10.1111/gcb.12627; pmid: 24865818
54. A. Pontes-Lopes *et al*., Drought-driven wildfire impacts on structure and dynamics in a wet Central Amazonian forest. *Proc. Biol. Sci.* **288**, 20210094 (2021). doi: 10.1098/rspb.2021.0094; pmid: 34004131
55. C. V. J. Silva *et al*., Estimating the multi-decadal carbon deficit of burned Amazonian forests. *Environ. Res. Lett.* **15**, 114023 (2020). doi: 10.1088/1748-9326/abb62c
56. J. Barlow, C. A. Peres, B. O. Lagan, T. Haugaasen, Large tree mortality and the decline of forest biomass following Amazonian wildfires. *Ecol. Lett.* **6**, 6–8 (2003). doi: 10.1046/j.1461-0248.2003.00394.x
57. A. C. Staver *et al*., Thinner bark increases sensitivity of wetter Amazonian tropical forests to fire. *Ecol. Lett.* **23**, 99–106 (2020). doi: 10.1111/ele.13409; pmid: 31642170
58. O. L. Phillips, S. Rose, A. M. Mendoza, P. N. Vargas, Resilience of southwestern Amazon forests to anthropogenic edge effects. *Conserv. Biol.* **20**, 1698–1710 (2006).
59. O. L. Phillips *et al*., Drought sensitivity of the Amazon rainforest. *Science* **323**, 1344–1347 (2009). doi: 10.1126/science.1164033; pmid: 19265020
60. C. V. J. Silva *et al*., Drought-induced Amazonian wildfires instigate a decadal-scale disruption of forest carbon dynamics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **373**, 20180043 (2018). doi: 10.1098/rstb.2018.0043; pmid: 30297477
61. D. Zuleta, A. Duque, D. Cardenas, H. C. Muller-Landau, S. J. Davies, Drought-induced mortality patterns and rapid biomass recovery in a terra firme forest in the Colombian Amazon. *Ecology* **98**, 2538–2546 (2017). doi: 10.1002/ecy.1950; pmid: 28719081
62. R. J. FletcherJr., Multiple edge effects and their implications in fragmented landscapes. *J. Anim. Ecol.* **74**, 342–352 (2005). doi: 10.1111/j.1365-2656.2005.00930.x
63. Brasil, "Quarta comunicação nacional do Brasil à Convenção-Quadro das Nações Unidas sobre Mudança do Clima" (Brasília, 2021).
64. A. Esquivel-Muelbert *et al*., Compositional response of Amazon forests to climate change. *Glob. Change Biol.* **25**, 39–56 (2019).
65. I. Numata, K. Khand, J. Kjaersgaard, M. A. Cochrane, S. S. Silva, Forest evapotranspiration dynamics over a fragmented forest landscape under drought in southwestern Amazonia. *Agric. For. Meteorol.* **306**, 108446 (2021). doi: 10.1016/j.agrformet.2021.108446
66. P. M. Brando *et al*., Droughts, Wildfires, and Forest Carbon Cycling: A Pantropical Synthesis. *Annu. Rev. Earth Planet. Sci.* **47**, 555–581 (2019). doi: 10.1146/annurev-earth-082517-010235
67. L. T. Smith, L. E. O. C. Aragão, C. E. Sabel, T. Nakaya, Drought impacts on children's respiratory health in the Brazilian Amazon. *Sci. Rep.* **4**, 3726 (2014). doi: 10.1038/srep03726; pmid: 24430803
68. N. de Magalhães, H. Evangelista, T. Condom, A. Rabatel, P. Ginot, Amazonian Biomass Burning Enhances Tropical Andean Glaciers Melting. *Sci. Rep.* **9**, 16914 (2019). doi: 10.1038/s41598-019-53284-1; pmid: 31780676
69. F. França, J. Barlow, B. Araújo, J. Louzada, Does selective logging stress tropical forest invertebrates? Using fat stores to examine sublethal responses in dung beetles. *Ecol. Evol.* **6**, 8526–8533 (2016). doi: 10.1002/ece3.2488; pmid: 28031804
70. A. C. Lees, C. A. Peres, Rapid avifaunal collapse along the Amazonian deforestation frontier. *Biol. Conserv.* **133**, 198–211 (2006). doi: 10.1016/j.biocon.2006.06.005
71. L. L. Jacob, B. S. Prudente, L. F. A. Montag, R. R. Silva, The effect of different logging regimes on the ecomorphological structure of stream fish assemblages in the Brazilian Amazon. *Hydrobiologia* **848**, 1027–1039 (2021). doi: 10.1007/s10750-020-04508-3
72. F. M. França *et al*., El Niño impacts on human-modified tropical forests: Consequences for dung beetle diversity and associated ecological processes. *Biotropica* **52**, 252–262 (2020). doi: 10.1111/btp.12756
73. V. Jirinec *et al*., Morphological consequences of climate change for resident birds in intact Amazonian rainforest. *Sci. Adv.* **7**, eabk1743 (2021). doi: 10.1126/sciadv.abk1743; pmid: 34767440
74. L. A. M. Mestre, M. A. Cochrane, J. Barlow, Long-term Changes in Bird Communities after Wildfires in the Central Brazilian Amazon. *Biotropica* **45**, 480–488 (2013). doi: 10.1111/btp.12026
75. L. G. Ziccardi, P. M. L. de Alencastro Graça, E. O. Figueiredo, P. M. Fearnside, Decline of large-diameter trees in a bamboo-dominated forest following anthropogenic disturbances in southwestern Amazonia. *Ann. For. Sci.* **76**, 110 (2019). doi: 10.1007/s13595-019-0901-4
76. C. J. Gardner, J. E. Bicknell, W. Baldwin-Cantello, M. J. Struebig, Z. G. Davies, Quantifying the impacts of defaunation on natural forest regeneration in a global meta-analysis. *Nat. Commun.* **10**, 4590 (2019). doi: 10.1038/s41467-019-12539-1; pmid: 31611554
77. J. Barlow, C. A. Peres, Effects of Single and Recurrent Wildfires on Fruit Production and Large Vertebrate Abundance in a Central Amazonian Forest. *Biodivers. Conserv.* **15**, 985–1012 (2006). doi: 10.1007/s10531-004-3952-1
78. L. N. Paolucci *et al*., Lowland tapirs facilitate seed dispersal in degraded Amazonian forests. *Biotropica* **51**, 245–252 (2019). doi: 10.1111/btp.12627
79. D. Vanham *et al*., Environmental footprint family to address local to planetary sustainability and deliver on the SDGs. *Sci. Total Environ.* **693**, 133642 (2019). doi: 10.1016/j.scitotenv.2019.133642; pmid: 31635013
80. J. Barlow *et al*., The future of hyperdiverse tropical ecosystems. *Nature* **559**, 517–526 (2018). doi: 10.1038/s41586-018-0301-1; pmid: 30046075
81. V. Galaz, B. Crona, A. Dauriach, B. Scholtens, W. Steffen, Finance and the Earth system – Exploring the links between financial actors and non-linear changes in the climate system. *Glob. Environ. Change* **53**, 296–302 (2018). doi: 10.1016/j.gloenvcha.2018.09.008
82. R. Rajão *et al*., The rotten apples of Brazil's agribusiness. *Science* **369**, 246–248 (2020). doi: 10.1126/science.aba6646; pmid: 32675358
83. A. T. Leite-Filho, B. S. Soares-Filho, J. L. Davis, G. M. Abrahão, J. Börner, Deforestation reduces rainfall and agricultural revenues in the Brazilian Amazon. *Nat. Commun.* **12**, 2591 (2021). doi: 10.1038/s41467-021-22840-7; pmid: 33972530
84. M. G. Ceddia, U. Gunter, A. Corriveau-Bourque, Land tenure and agricultural expansion in Latin America: The role of Indigenous Peoples' and local communities' forest rights. *Glob. Environ. Change* **35**, 316–322 (2015). doi: 10.1016/j.gloenvcha.2015.09.010
85. M. Lentini, R. Palmieri, L. Sobral, "Atividade florestal madeireira no norte do Pará: do projeto Jari ao combate à ilegalidade e ao advento das concessões florestais" in *Florestas de Valor. Desenvolvimento territorial e sustentabilidade: a experiência do IMAFLORA no Norte do Pará*, R. Palmieri, Ed. (IMAFLORA, Belém, 2020), pp. 64–93.
86. C. Adams, R. S. S. Murrieta, R. A. Sanches, Agricultura e alimentação em populações ribeirinhas das várzeas do Amazonas: Novas perspectivas. *Ambiente Soc.* **8**, 65–86 (2005). doi: 10.1590/S1414-753X2005000100005
87. S. L. Francelino-Gonçalves-Dias, P. Mendonça, "Deforestation and Slave Labour in the Amazon: Contesting the sustainability of the cattle industry" in *7th International Critical Management Studies (CMS) Conference* (Naples, 2011).
88. R. D. Garrett *et al*., Explaining the persistence of low income and environmentally degrading land uses in the Brazilian Amazon. *Ecol. Soc.* **22**, art27 (2017). doi: 10.5751/ES-09364-220327
89. P. Shanley, M. da Serra Silva, T. Melo, R. Carmenta, R. Nasi, From conflict of use to multiple use: Forest management innovations by small holders in Amazonian logging frontiers. *For. Ecol. Manage.* **268**, 70–80 (2012). doi: 10.1016/j.foreco.2011.05.041
90. V. L. Camilotti, P. Pinho, E. S. Brondízio, M. I. S. Escada, The Importance of Forest Extractive Resources for Income Generation and Subsistence among Caboclos and Colonists in the Brazilian Amazon. *Hum. Ecol. Interdiscip. J.* **48**, 17–31 (2020). doi: 10.1007/s10745-020-00127-7
91. S. C. das C. Xavier *et al*., Lower Richness of Small Wild Mammal Species and Chagas Disease Risk. *PLOS Negl. Trop. Dis.* **6**, e1647 (2012).
92. F. Michalski, C. A. Peres, Disturbance-mediated mammal persistence and abundance-area relationships in Amazonian forest fragments. *Conserv. Biol.* **21**, 1626–1640 (2007). doi: 10.1111/j.1523-1739.2007.00797.x; pmid: 18173486
93. P. F. Pinho *et al*., Ecosystem protection and poverty alleviation in the tropics: Perspective from a historical evolution of policy-making in the Brazilian Amazon. *Ecosyst. Serv.* **8**, 97–109 (2014). doi: 10.1016/j.ecoser.2014.03.002
94. A. McGregor, S. Coulthard, L. Camfield, "Measuring what matters: the role of well-being methods in development policy and practice", ODI Development Progress (London, 2015), 20-p. Available at: https://cdn.odi.org/media/documents/9688.pdf.
95. A. Y. Vittor, G. Z. Laporta, M. A. M. Sallum, R. T. Walker, The COVID-19 crisis and Amazonia's indigenous people: Implications for conservation and global health. *World Dev.* **145**, 105533 (2021). doi: 10.1016/j.worlddev.2021.105533; pmid: 36570383
96. W. A. Campanharo, A. P. Lopes, L. O. Anderson, T. F. M. R. da Silva, L. E. O. C. Aragão, Translating Fire Impacts in Southwestern Amazonia into Economic Costs. *Remote Sens. (Basel)* **11**, 764 (2019). doi: 10.3390/rs11070764
97. M. O. Nawaz, D. K. Henze, Premature Deaths in Brazil Associated With Long-Term Exposure to $PM_{2.5}$ From Amazon Fires Between 2016 and 2019. *Geohealth* **4**, GH000268 (2020). doi: 10.1029/2020GH000268; pmid: 32864540
98. Human Rights Watch, "Rainforest mafias: How violence and impunity fuels deforestation in Brazil's Amazon," Human Rights Watch, New York (2019).
99. A. B. Chimeli, R. R. Soares, The Use of Violence in Illegal Markets: Evidence from Mahogany Trade in the Brazilian Amazon. *Am. Econ. J. Appl. Econ.* **9**, 30–57 (2017). doi: 10.1257/app.20160055
100. K. Whyte, Too late for indigenous climate justice: Ecological and relational tipping points. *Wiley Interdiscip. Rev. Clim. Change* **11**, e603 (2020). doi: 10.1002/wcc.603
101. R. Carmenta, E. Coudel, A. M. Steward, Forbidden fire: Does criminalising fire hinder conservation efforts in swidden landscapes of the Brazilian Amazon? *Geogr. J.* **185**, 23–37 (2019). doi: 10.1111/geoj.12255
102. P. Daszak *et al*., "Workshop Report on Biodiversity and Pandemics of the Intergovernmental Platform on Biodiversity and Ecosystem Services" (Bonn, 2020); https://doi.org/10.5281/zenodo.4147317.
103. J. H. Ellwanger *et al*., Beyond diversity loss and climate change: Impacts of Amazon deforestation on infectious diseases and public health. *An Acad Bras Cienc.* **92** (2020). doi: 10.1590/0001-3765202020191375
104. J. A. Foley *et al*., Amazonia revealed: Forest degradation and loss of ecosystem goods and services in the Amazon Basin. *Front. Ecol. Environ.* **5**, 25–32 (2007). doi: 10.1890/1540-9295(2007)5[25:ARFDAL]2.0.CO;2
105. A. P. D. Aguiar, G. Câmara, M. I. S. Escada, Spatial statistical analysis of land-use determinants in the Brazilian Amazonia: Exploring intra-regional heterogeneity. *Ecol. Modell.* **209**, 169–188 (2007). doi: 10.1016/j.ecolmodel.2007.06.019
106. D. M. Lapola *et al*., Impacts of climate change and the end of deforestation on land use in the Brazilian Legal Amazon. *Earth Interact.* **15**, 1–29 (2011). doi: 10.1175/2010EI333.1
107. W. F. Laurance *et al*., Environment. The future of the Brazilian Amazon. *Science* **291**, 438–439 (2001). doi: 10.1126/science.291.5503.438; pmid: 11228139
108. B. S. Soares-Filho *et al*., Modelling conservation in the Amazon basin. *Nature* **440**, 520–523 (2006). doi: 10.1038/nature04389; pmid: 16554817

109. Y. Le Page *et al*., Synergy between land use and climate change increases future fire risk in Amazon forests. *Earth Syst. Dyn.* **8**, 1237–1246 (2017). doi: 10.5194/esd-8-1237-2017
110. D. Armenteras, U. Murcia, T. M. González, O. J. Barón, J. E. Arias, Scenarios of land use and land cover change for NW Amazonia: Impact on forest intactness. *Glob. Ecol. Conserv.* **17**, e00567 (2019). doi: 10.1016/j.gecco.2019.e00567
111. B. L. De Faria *et al*., Current and future patterns of fire-induced forest degradation in Amazonia. *Environ. Res. Lett.* **12**, 095005 (2017). doi: 10.1088/1748-9326/aa69ce
112. T. O. Assis, A. P. D. Aguiar, C. von Randow, C. A. Nobre, Projections of future forest degradation and CO_2 emissions for the Brazilian Amazon. *Sci. Adv.* **8**, eabj3309 (2022). doi: 10.1126/sciadv.abj3309; pmid: 35704589
113. A. P. D. Aguiar *et al*., Land use change emission scenarios: Anticipating a forest transition process in the Brazilian Amazon. *Glob. Chang. Biol.* **22**, 1821–1840 (2016). doi: 10.1111/gcb.13134; pmid: 26511401
114. R. B. de Andrade *et al*., Scenarios in tropical forest degradation: Carbon stock trajectories for REDD. *Carbon Balance Manag.* **12**, 6 (2017). doi: 10.1186/s13021-017-0074-0; pmid: 28413850
115. L. E. O. C. Aragão, Y. E. Shimabukuro, The incidence of fire in Amazonian forests with implications for REDD. *Science* **328**, 1275–1278 (2010). doi: 10.1126/science.1186925; pmid: 20522775
116. P. Delacote, A. Angelsen, Reducing Deforestation and Forest Degradation: Leakage or Synergy? *Land Econ.* **91**, 501–515 (2015). doi: 10.3368/le.91.3.501
117. S. N. Panfil, C. A. Harvey, REDD+ and Biodiversity Conservation: A Review of the Biodiversity Goals, Monitoring Methods, and Impacts of 80 REDD+ Projects. *Conserv. Lett.* **9**, 143–150 (2016). doi: 10.1111/conl.12188
118. M. G. Fonseca *et al*., Effects of climate and land-use change scenarios on fire probability during the 21st century in the Brazilian Amazon. *Glob. Chang. Biol.* **25**, 2931–2946 (2019). doi: 10.1111/gcb.14709; pmid: 31304669
119. T. E. Lovejoy, C. Nobre, Amazon Tipping Point. *Sci. Adv.* **4**, eaat2340 (2018). doi: 10.1126/sciadv.aat2340; pmid: 29492460
120. U. Pascual *et al*., Governing for Transformative Change across the Biodiversity–Climate–Society Nexus. *Bioscience* **72**, 684–704 (2022). doi: 10.1093/biosci/biac031
121. J. Reed, J. Van Vianen, E. L. Deakin, J. Barlow, T. Sunderland, Integrated landscape approaches to managing social and environmental issues in the tropics: Learning from the past to guide the future. *Glob. Chang. Biol.* **22**, 2540–2554 (2016). doi: 10.1111/gcb.13284; pmid: 26990574
122. M. Skutsch, E. Turnhout, REDD+: If communities are the solution, what is the problem? *World Dev.* **130**, 104942 (2020). doi: 10.1016/j.worlddev.2020.104942
123. C. Streck, REDD+ and leakage: Debunking myths and promoting integrated solutions. *Clim. Policy* **21**, 843–852 (2021). doi: 10.1080/14693062.2021.1920363
124. R. DeFries *et al*., Improved household living standards can restore dry tropical forests. *Biotropica* **54**, 1480–1490 (2022). doi: 10.1111/btp.12978; pmid: 36582545
125. H. K. Gibbs *et al*., Brazil's Soy Moratorium. *Science* **347**, 377–378 (2015). doi: 10.1126/science.aaa0181; pmid: 25613879
126. D. F. Amaral, J. B. de Souza Ferreira Filho, A. L. S. Chagas, M. Adami, Expansion of soybean farming into deforested areas in the amazon biome: The role and impact of the soy moratorium. *Sustain. Sci.* **16**, 1295–1312 (2021). doi: 10.1007/s11625-021-00942-x
127. S. Hancock, C. McGrath, C. Lowe, I. Davenport, I. Woodhouse, Requirements for a global lidar system: Spaceborne lidar with wall-to-wall coverage. *R. Soc. Open Sci.* **8**, 211166 (2021). doi: 10.1098/rsos.211166; pmid: 34877004
128. S. Francini, G. D'Amico, E. Vangi, C. Borghi, G. Chirici, Integrating GEDI and Landsat: Spaceborne Lidar and Four Decades of Optical Imagery for the Analysis of Forest Disturbances and Biomass Changes in Italy. *Sensors* **22**, 2015 (2022). doi: 10.3390/s22052015; pmid: 35271161
129. C. M. Souza Jr. *et al*., Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and Earth Engine. *Remote Sens.* **12** (2020). doi: 10.3390/rs12172735
130. S. Saatchi *et al*., Persistent effects of a severe drought on Amazonian forest canopy. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 565–570 (2013). doi: 10.1073/pnas.1204651110; pmid: 23267086

ACKNOWLEDGMENTS

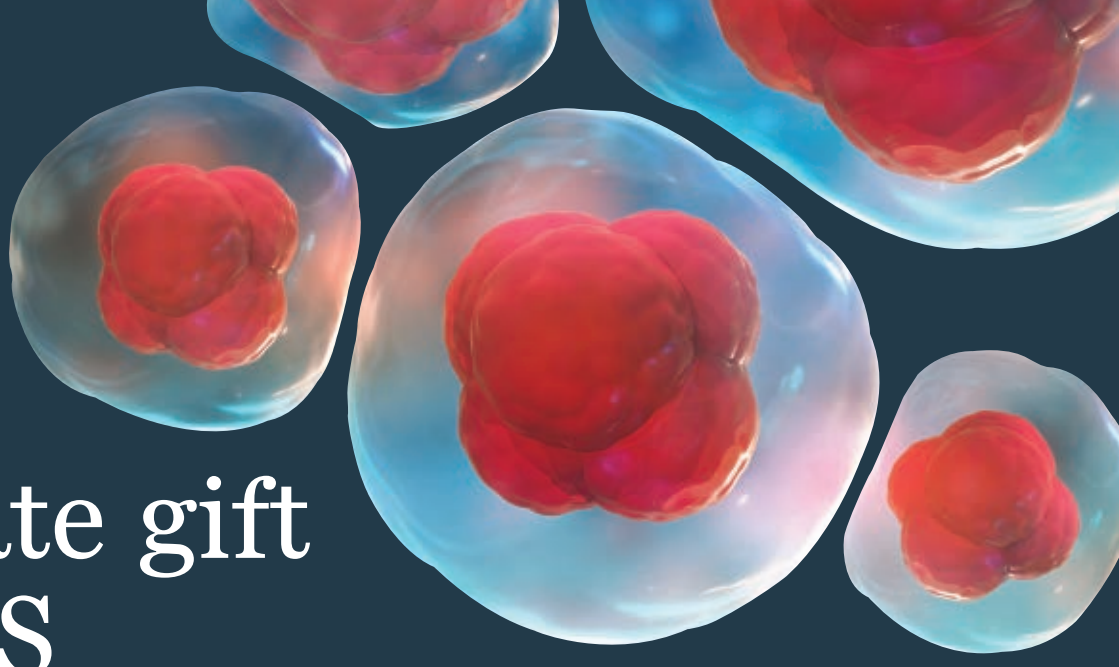
This Review was supported by the AIMES (Analysis, Integration and Modeling of the Earth System) Project through the workshop "Degradation of tropical forests: Observations, modeling and socio-environmental implications" held 11 to 13 November 2019 in Manaus, Brazil. D.M.L. was supported through São Paulo Research Foundation – FAPESP grants 2015/02537-7 and 2020/08940-6 and CNPq (309074/2021-5). J.B. and E.B. were funded by the UK Natural Environment Research Council (NE/S01084X/1). E.B., J.B., and J.F. are grateful to CNPq 441949/2018-5 (Sem-Flama) and 441573/2020-7 (PELD-RAS), as well as to the BNP-Paribas Bioclimate grant. J.F. is also grateful to CNPq 314242/2021-0. L.E.O.C.A. thanks CNPq (314416/2020-0) and the Brazilian Space Agency (AEB). C.H.L.S.-J. was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) grant 001 and by the University of Manchester through the project entitled "Forest fragmentation mapping of Amazon and its vulnerable margin Amazon-Cerrado transition forests." Part of this work was carried out at the Jet Propulsion Laboratory under a contract with NASA. L.O.A. was supported by FAPESP (2020/08916-8, 2016/02018-2, 2020/16457-3, 2020/15230-5) and CNPq (441949/2018-5, 409531/2021-9, 314473/2020-3). L.G. was supported by FAPESP (2016/02018-2). C.v.R. thanks CNPq (314780/2020-3). V.H.G.-V. was supported by the National Aeronautics and Space Administration through the A.50 Group on Earth Observations Work Program [grant no. 80NSSC18K0339]. P.M.F. was supported by CNPq (312450/2021-4) and FAPESP (2020/08916-8)/Amazonas Research Foundation-FAPEAM (01.02.016301.00289/2021). A.C.R. was supported by the NASA Earth Science Division Climate Impacts Group Funding (509496.02.80.01.03). We are grateful to G. G. Perez for assistance with data deposition, A. Argozino for producing the summary figure and Figs. 1 and 4, and S. Garcia and C. A. Quesada for comments on earlier versions of this manuscript. **Author contributions:** D.M.L., P.P., J.B., L.E.O.C.A., E.B., R.C., H.M.L., H.S., C.V.J.S., and C.H.L.S.-J. were responsible for the conceptualization of this Review, data visualization and analysis, and writing of the original draft. H.M.L. and D.M.L. acquired funding that made this study possible. All other authors contributed additional ideas and revised subsequent drafts. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All previously published data shown here can be obtained through the references cited in the supplementary materials (*43*). The codes and data resulting from the analyses of such previously published material are freely available at https://doi.org/10.25824/redu/EGJAYI. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abp8622
Materials and Methods
Tables S1 and S2

Submitted 10 March 2022; accepted 28 November 2022
10.1126/science.abp8622

1848 SOCIETY
AAAS



An estate gift to AAAS

Going all the way back to 1848, our founding year, the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

Today, we are dedicated to advocating for science and scientific evidence to be fully and positively integrated into public policy and for the community to speak with one voice to advance science and engineering in the United States and around the world.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you will become a member of our 1848 Society and will help fuel our work on behalf of science and society – including publishing the world's most promising, innovative research in the *Science* family of journals and engaging in the issues that matter locally, nationally and around the world.

"As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all."

—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

Legal Name: American Association for the Advancement of Science
Federal Tax ID Number: 53-0196568
Address: 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at **www.aaas.org/1848Society**.

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

cut here

Yes, I would like more information about joining the AAAS 1848 Society.

1848 SOCIETY
AAAS

PLEASE CONTACT ME AT:

Name: _____

Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____

RETURN THIS FORM TO:
AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

RESEARCH ARTICLE SUMMARY

VACCINES

Low protease activity in B cell follicles promotes retention of intact antigens after immunization

Aereas Aung, Ang Cui, Laura Maiorino, Ava P. Amini, Justin R. Gregory, Maurice Bukenya, Yiming Zhang, Heya Lee, Christopher A. Cottrell, Duncan M. Morgan, Murillo Silva, Heikyung Suh, Jesse D. Kirkpatrick, Parastoo Amlashi, Tanaka Remba, Leah M. Froehle, Shuhao Xiao, Wuhbet Abraham, Josetta Adams, J. Christopher Love, Phillip Huyett, Douglas S. Kwon, Nir Hacohen, William R. Schief, Sangeeta N. Bhatia, Darrell J. Irvine*

INTRODUCTION: Antigens in vaccines are designed to induce antibodies that bind to key epitopes on the surfaces of pathogens, promoting microbe elimination. Despite advances in the engineering of immunogens that closely mimic native pathogen protein structures, not all vaccines elicit protective humoral immunity. We hypothesized that the integrity of vaccine antigens in vivo could be an important factor. Extracellular antigen degradation in lymph nodes (LNs) could limit the generation of protective antibody responses and also create competing irrelevant responses against antigen breakdown products. However, vaccine antigen stability in vivo is poorly understood.

RATIONALE: Extracellular protease activity within the LN has not been studied, and how such potential proteolytic activity could affect immunogen structure and subsequent immune responses is unknown. We used HIV immunogens as a model system and measured antigen stability in the LN, identified degradative proteases, and explored vaccination strategies to maximize the delivery of intact antigens to B cells.

RESULTS: We investigated antigen stability within the LN by conjugating antigens with dyes that undergo a loss in fluorescence resonance energy transfer (FRET) upon structural breakdown. After vaccination, antigen was rapidly degraded in the subcapsular sinus (SCS) and extrafollicular regions of the LN within 48 hours, with approximately half of the antigen breakdown occurring extracellularly. By contrast, antigen localized to B cell follicles remained intact.

To determine the cause of antigen degradation, we examined protease expression and activity in LNs. RNA sequencing, histology, and imaging zymography revealed that extracellular proteases and protease activity were present at high levels in the SCS and extrafollicular regions, but were present at low levels within follicles. The identified metalloproteases degraded antigen in vitro, and their inhibition in vivo increased intact antigen levels within the LN.

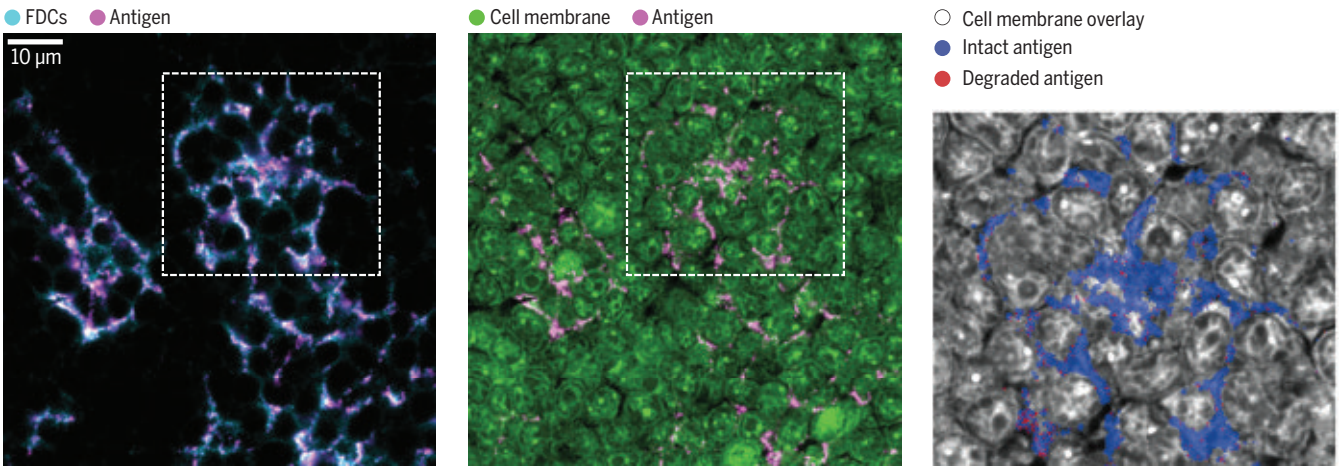
Although antigen captured by follicular dendritic cells (FDCs) has been reported to be retained in an intact state, the mechanisms for this preservation have remained unclear. We tested the importance of low follicular protease activity by adoptively transferring protease-expressing polyclonal B cells into the LNs of immunized mice. We observed a significant decrease in FDC-captured antigen stability, indicating that low follicular protease activity is important for intact antigen retention.

Motivated by these findings, we tested vaccination strategies to maximize the humoral response to intact antigen. Immunization by "extended dosing" regimens or nanoparticle formulations that rapidly targeted antigen to FDCs were compared with traditional bolus-soluble antigen vaccination (in which little follicular uptake of antigen is observed). FDC-targeting vaccinations led to large germinal centers (GCs), with high numbers of B cells recognizing intact antigen. By contrast, traditional bolus immunization resulted in small GCs with B cells that equally recognized intact antigen and breakdown products. Similar patterns of response were observed in serum antibody titers.

CONCLUSION: Here, we have shown that rapid antigen degradation can occur in LNs after vaccination except within follicles. This degradation is mediated by spatially compartmentalized proteolysis within the tissue. Vaccine strategies promoting rapid follicular antigen localization strongly promoted antibody generation against intact antigen without amplifying responses to irrelevant antigen breakdown products. Such approaches may enable more effective vaccines against difficult pathogens such as HIV. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: djirvine@mit.edu
Cite this article as A. Aung *et al.*, *Science* **379**, eabn8934 (2023). DOI: 10.1126/science.abn8934

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abn8934



Stability of FDC-captured antigen after immunization. Left and middle: HIV antigens (magenta) localized to FDCs (cyan) that reside among cells in the follicle (green) 3 days after immunization. Right: FRET analysis of antigen structural integrity within the outlined region shows that antigen localized to FDCs remains intact. Scale bar, 10 μm.

RESEARCH ARTICLE

VACCINES

Low protease activity in B cell follicles promotes retention of intact antigens after immunization

Aereas Aung[1], Ang Cui[2,3], Laura Maiorino[1], Ava P. Amini[1,3,4,5], Justin R. Gregory[1], Maurice Bukenya[1]†, Yiming Zhang[1,6], Heya Lee[1], Christopher A. Cottrell[7,8], Duncan M. Morgan[1,9], Murillo Silva[1], Heikyung Suh[1], Jesse D. Kirkpatrick[1,3], Parastoo Amlashi[1], Tanaka Remba[1], Leah M. Froehle[10], Shuhao Xiao[1], Wuhbet Abraham[1], Josetta Adams[1], J. Christopher Love[1,2,9,10], Phillip Huyett[11], Douglas S. Kwon[10], Nir Hacohen[2,12], William R. Schief[8,10,13,14], Sangeeta N. Bhatia[1,2,3,15,16,17,18], Darrell J. Irvine[1,6,8,10,18,19]*

The structural integrity of vaccine antigens is critical to the generation of protective antibody responses, but the impact of protease activity on vaccination in vivo is poorly understood. We characterized protease activity in lymph nodes and found that antigens were rapidly degraded in the subcapsular sinus, paracortex, and interfollicular regions, whereas low protease activity and antigen degradation rates were detected in the vicinity of follicular dendritic cells (FDCs). Correlated with these findings, immunization regimens designed to target antigen to FDCs led to germinal centers dominantly targeting intact antigen, whereas traditional immunizations led to much weaker responses that equally targeted the intact immunogen and antigen breakdown products. Thus, spatially compartmentalized antigen proteolysis affects humoral immunity and can be exploited.

After vaccination, humoral immune responses begin by B cells binding with their receptors to cognate antigen, followed by the formation of germinal centers (GCs), where these cells undergo proliferation and affinity maturation, leading to the production of high-affinity antibodies against the target antigen (*1*–*3*). Factors determining the makeup of the eventual affinity-matured polyclonal antibody response remain incompletely understood, although the precursor frequency of antigen-specific B cells, the affinity of precursors for the antigen, antigen complexity, follicular helper T cell–derived signals, and antibody feedback all contribute (*4*–*8*). In addition, the duration of antigen exposure and the amount of antigen available to B cells plays an important role (*9*, *10*).

We hypothesized that the structural integrity of the antigen in vivo is an additional important factor. To elicit protective responses, antigens need to present neutralizing epitopes that are faithful structural mimics of the target pathogen, which are often complex three-dimensional surfaces (*11*). Disruption of these epitopes could not only limit the activation of B cells with the capacity to produce neutralizing antibodies, but might also create distracting de novo epitopes irrelevant for protective immunity. It has been reported that model protein antigens can be rapidly proteolyzed as they reach the subcapsular sinus (SCS) of lymph nodes (LNs), and this antigen cleavage was linked to protease activity in serum and interstitial fluid (*12*). Such pathways of antigen breakdown might at least partially explain the substantial proportion of B cells that enter GC reactions but do not detectably bind to the immunizing antigen (*7*, *9*).

By contrast, several lines of evidence suggest that antigen trapped on dendrites of follicular dendritic cells (FDCs) may remain intact over extended time periods. Early studies showed that FDC-bound antigen recovered from LNs after 12 weeks could be recognized by epitope-sensitive monoclonal antibodies (mAbs) and were eluted in size-exclusion chromatography in a manner suggesting gross antigen integrity (*13*). HIV virions deposited on FDCs in mice can be extracted from LNs and functional viral particles recovered over several months, although the quantitative proportion of particles that are infective is not clear (*14*). FDCs have also been shown to cyclically internalize and recycle trapped antigen, which may protect it from extracellular degradation (*15*). These data collectively suggest that the follicles, and the FDC networks in particular, may be sites within LNs where antigens are protected from degradation, whereas regions such as the sinuses may be areas of high proteolytic activity. To our knowledge, however, the nature of protease activity in lymphoid organs has not been studied, and how antigen proteolysis affects the immune response to vaccines is poorly understood.

To shed light on the fate of antigens during the primary immune response, we developed a FRET-based approach to track the integrity of antigens after subunit vaccine immunization and analyzed the spatial pattern of protease expression and activity in LNs. Unexpectedly, we found a pronounced spatial variation in protease activity, with high levels of antigen breakdown and protease expression in extrafollicular regions of mouse and human lymphoid tissues, but low levels of protease activity and high retention of antigen integrity over time within the FDC network of B cell follicles. Prompted by these findings, we evaluated the impact of antigen localization on the specificity of GC B cell responses, and found evidence that FDC-targeted protein immunizations achieve substantially greater proportions of antigen-specific B cell responses targeting conformationally intact epitopes compared with traditional bolus vaccination.

RESULTS

Monitoring antigen integrity using FRET analysis

To investigate vaccine antigen stability after immunization, we labeled immunogens with paired small-molecule dyes capable of undergoing fluorescence resonance energy transfer (FRET) to detect gross disruptions of antigen structure in situ in tissues. We hypothesized that antigen proteolysis would lead to the separation of FRET donor and acceptor dyes, lowering FRET signals in proportion to the degree of antigen degradation, a process that can be tracked by microscopy or flow cytometry (Fig. 1A). For microscopy-based analysis, we used the acceptor photobleaching method (*16*), in which the emission of a donor dye (Cy3) is measured before and after the bleaching of an acceptor dye (Cy5) to monitor antigen integrity (fig. S1A). This technique avoids complications of donor and acceptor excitation and emission cross-talk because only the donor emission is analyzed, and it is independent of dye concentrations and ratios. As a test case, we first focused on the clinical vaccine candidate eOD-GT8 60mer (eOD-60mer), an

[1]Koch Institute for Integrative Cancer Research, MIT, Cambridge, MA, USA. [2]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [3]Harvard-MIT Division of Health Sciences and Technology, MIT, Cambridge, MA, USA. [4]Program in Biophysics, Harvard University, Boston, MA, USA. [5]Microsoft Research New England, Cambridge, MA, USA. [6]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. [7]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA. [8]Consortium for HIV/AIDS Vaccine Development, The Scripps Research Institute, La Jolla, CA, USA. [9]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. [10]Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA. [11]Department of Otolaryngology-Head and Neck Surgery, Massachusetts Eye and Ear Infirmary and Harvard Medical School, Boston, MA, USA. [12]Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [13]Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, CA, USA. [14]IAVI Neutralizing Antibody Center, The Scripps Research Institute, La Jolla, CA, USA. [15]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA. [16]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [17]Wyss Institute at Harvard, Boston, MA, USA. [18]Howard Hughes Medical Institute, Cambridge, MA, USA. [19]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA.
*Corresponding author. Email: djirvine@mit.edu
†Present address: Department of Immunology, Duke University, Durham, NC, USA.

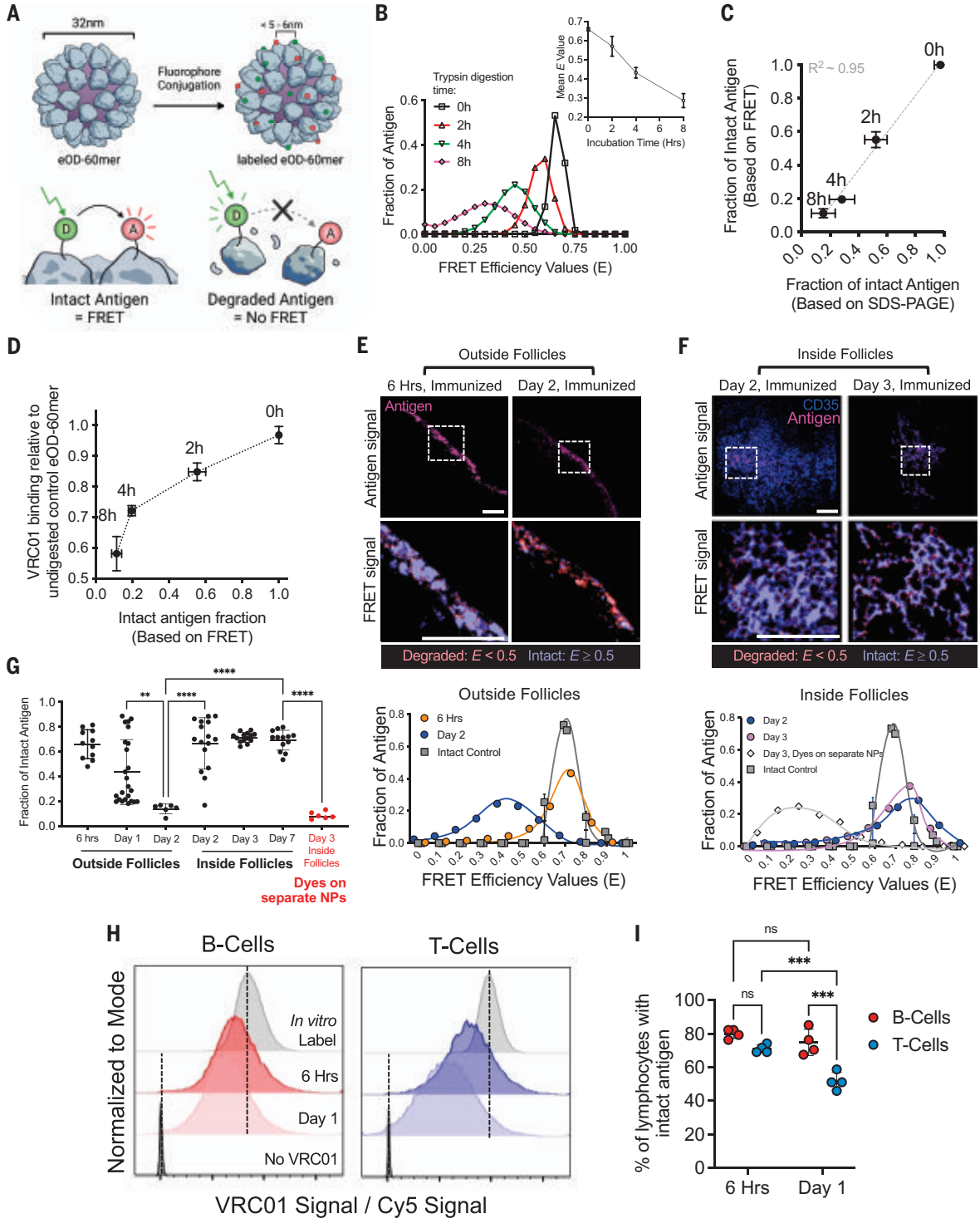


Fig. 1. FRET-based analysis of antigen integrity reveals spatially compartmentalized antigen degradation in LNs. (**A**) Schematic of eOD-60mer labeled with donor and acceptor dyes to track antigen integrity by FRET. (**B** to **D**) FRET analysis of eOD-60mer$_{40}$ integrity after trypsin digestion (n = 3 samples/time point). Shown are representative E histograms and mean E values (B), proportion of intact eOD-60mer$_{40}$ based on FRET versus SDS-PAGE (C), and VRC01 mAb binding to digested eOD6-mer$_{40}$ versus FRET-measured proportion of intact eOD60-mer$_{40}$ (D). (**E** to **G**) C57BL/6 mice (n = 4 animals/group) were immunized with 10 μg of eOD-60mer$_{40}$ and 5 μg of saponin adjuvant. At the indicated time points, LNs were flash frozen and sectioned. (E and F) Representative images of eOD-60mer$_{40}$ within the SCS ["Outside Follicles" (E)] versus B cell follicles ["Inside Follicles" (F)]. Top images show antigen and CD35 staining; bottom images show intact or degraded antigens in the dashed boxes as false colors at higher magnification. Corresponding E histograms within dashed boxes are shown compared with intact eOD-60mer$_{40}$ coated on coverslips. "Day 3, Dyes on separate NPs" in (F) indicates mice immunized with saponin adjuvant and 5 μg of two different eOD-60mers, each labeled with only Cy3 or Cy5 dyes. Scale bars, 100 μm. (G) Mean fraction of intact eOD-60mer$_{40}$ at the indicated LN locations at different time points after immunization. Each point represents one region from one tissue section.

Data were collected from at least eight tissue sections from eight LNs. *P* values were determined by one-way ANOVA with Tukey's post test. (**H** and **I**) C57BL/6 mice (n = 4 animals/group) were immunized with 20 μg of anti-CD45 mAb conjugated to Cy5-labeled eOD monomer and 5 μg of saponin adjuvant. LN cells were isolated at 6 hours or 1 day after immunization, stained with VRC01 mAb, and analyzed by flow cytometry. (H) Histograms of VRC01 binding normalized to eOD monomer-Cy5 signal showing gating on T cells or B cells. Vertical dashed lines indicate the mean fluorescence intensities of intact eOD-anti-CD45 freshly bound to splenocytes in vitro (positive control) versus cells analyzed in the absence of VRC01 (negative control). (I) Percentage of lymphocytes with intact eOD monomer as determined by VRC01 binding. *P* values were determined by two-way ANOVA with Tukey's post test. All plots show mean ± SD.

~30-nm-diameter protein nanoparticle (NP) presenting 60 copies of an HIV Env gp120-engineered outer domain (Fig. 1A). eOD-60mer is used as a priming immunogen to initiate CD4 binding site–directed broadly neutralizing antibodies against HIV (*17*–*21*). This NP antigen accommodated labeling with at least 40 dyes per particle (~20 Cy3 and ~20 Cy5 dyes, eOD-60mer$_{40}$) without disrupting the binding of the broadly neutralizing antibody VRC01 to the CD4-binding site epitope presented by the eOD-GT8 immunogen (fig. S1B). This degree of labeling represents modification of only 3.5% of the total amino acids of the particle immunogen. We previously showed that after a primary immunization, mannose-binding lectin (MBL) present in serum binds to glycans densely displayed on the surface of the eOD-60mer, leading to subsequent complement deposition on the particles and trafficking of the antigen to FDCs in the LN (*22*). Complement binding to unmodified or dye-labeled eOD-60mer$_{40}$ incubated in vitro in mouse serum was not statistically different (fig. S1C). eOD-60mer$_{40}$ coated onto glass coverslips showed a high level of FRET on excitation at 555 nm, but upon photobleaching of the Cy5 dye, emission in the Cy3 donor channel increased by twofold (fig. S1D). Such a change in Cy3 fluorescence was not observed for NPs labeled with only Cy3 or only Cy5, or when Cy3-labeled NPs were mixed with Cy5-labeled NPs and coated on coverslips together (fig. S1, E to H). These changes in acceptor emission before and after photobleaching can be quantified to determine the mean FRET efficiency (E) (see the materials and methods). Of importance for our subsequent application to frozen tissue sections, FRET signals of intact eOD-60mer$_{40}$ as a fresh protein or after flash freezing were not statistically different (fig. S1, I and J). On the basis of these findings, we focused on NPs labeled with ~40 dyes total for further studies.

To determine how E relates to the structural integrity of the particles, we calculated E distributions of intact or trypsin-digested eOD-60mer$_{40}$. Proteolyzed eOD-60mer$_{40}$ showed the development of breakdown products by gel electrophoresis (fig. S1K), which were accompanied by a shift in the distribution of measured E values for the particles to lower values with increasing digestion time (Fig. 1B). The fraction of fully intact antigen as determined by FRET (calculated as the proportion of material with an E value overlapping the undigested control particle E distribution) was highly correlated to the fraction of intact antigen as measured by SDS–polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 1C). VRC01 binding to eOD-60mer$_{40}$ captured on ELISA plates also decreased with increasing trypsin digestion time, demonstrating that the loss of FRET signal was not due simply to particle disassembly (Fig. 1D). Altogether, these data suggest that FRET signals from labeled 60mer NPs effectively report on the retention of intact immunogen structure.

Antigens are rapidly proteolyzed in the SCS and extrafollicular regions but protected in follicles

We next used this FRET-based approach to analyze antigen integrity in vivo in mice after immunization. We previously reported that eOD-60mer initially accumulates in the SCS of draining LNs, but over the course of several days, NPs are trafficked to the FDC network of follicles in an MBL- and complement-dependent manner, where they are retained for more than a week (*22*). To assess the structural integrity of the NPs during this process, we immunized mice with eOD-60mer$_{40}$ and a saponin-based adjuvant, and at various time points after immunization, draining LNs were extracted, flash frozen, and cryosectioned for acceptor photobleaching FRET analysis (fig. S2A). FRET efficiencies were converted to a measure of percent intact antigen by comparing E values measured in the tissue with that of control fresh antigen coated on glass coverslips. At early times (2 to 6 hours after immunization), eOD-60mer$_{40}$ accumulated in the SCS and could also be detected at very low levels in follicles or extrafollicular regions within the LN parenchyma (at roughly equal levels in these latter two locations; fig. S2, B and C). At later times, antigen continued to accumulate in the SCS, but over ~48 hours, antigen levels in the SCS decreased, whereas steadily increasing amounts were detected on FDCs (Fig. 1, E and F, and fig. S2, D and E). At very early times, most eOD-60mer$_{40}$ was intact regardless of its location (Fig. 1G). However, unexpectedly, we observed location-dependent patterns of antigen degradation developing over time. Nanoparticles localized in the sinus or interfollicular regions of the LN exhibited a substantial loss in E between 6 and 48 hours (Fig. 1E, "outside follicles"). By contrast, eOD-60mer$_{40}$ localized within the FDC network of follicles at the same time points exhibited high E values (Fig. 1F, "inside follicles"). Quantification of the proportion of intact NPs from many LNs revealed that most eOD-60mer$_{40}$ located outside of follicles was degraded within 48 hours, whereas 70 to 80% of the antigen localized to the FDC network remained intact through at least 7 days (Fig. 1G). Imaging of LNs from mice immunized with control mixtures of Cy3-eOD-60mer and Cy5-eOD-60mer showed negligible FRET signals for FDC-localized antigens, suggesting that the sustained FRET signals observed for FDC-localized antigen were not due to interparticle FRET (Fig. 1G, "dyes on separate NPs"). To determine whether similar spatially distinct patterns of antigen degradation are observed with a completely different antigen and nanoparticle carrier, we also imaged LNs after injection of FRET-labeled influenza hemagglutinin-ferritin nanoparticles (HA-NPs), representative of a class of NP vaccine in clinical trials (*23*). These NPs also accumulated on FDCs over time (*22*, *24*) (fig. S2, F and G), and like eOD-60mer, HA-NPs degraded rapidly in extrafollicular regions but were highly protected in follicles (fig. S2H).

As an orthogonal measure of antigen integrity, we stained LN sections with VRC01 antibody to detect 60mer with an intact CD4-binding site epitope, and this approach revealed similar results: eOD-60mer in the SCS showed a high level of VRC01 staining at 6 hours after immunization, but this staining was almost completely lost by 48 hours for material outside of the FDC network (fig. S3, A and B). By contrast, NP antigen localized to FDCs retained VRC01 staining even at 3 days after immunization (fig. S3, A and B). Loss of VRC01 staining in extrafollicular sites was not due to an early antibody response blocking the epitope because the same loss in VRC01 binding was observed in LNs from B cell receptor–transgenic MD4 mice, in which B cells express an antigen receptor specific for an irrelevant antigen (*25*) (fig. S3C).

It was unclear whether the retention of eOD-60mer integrity reflected a specific protective activity of FDCs or of the follicles in general. To answer this question, we devised an approach to target antigen into both follicles and the paracortex simultaneously. Antibodies to CD45 bind to lymphocytes and remain cell surface localized for multiple days (*26*). We exploited this biology to target eOD to the surfaces of T cells and B cells in LNs by conjugating dye-labeled monomeric eOD to an anti-CD45 mAb (fig. S4A). The antibody conjugate efficiently labeled CD45$^+$ LN cells in vitro, and a

flow cytometry–based fluorescence quenching assay (*27*) showed that >90% of the conjugate bound to the cells was extracellular (fig. S4, B to E). Mice were immunized with Cy5-labeled eOD-anti-CD45 and saponin adjuvant, and LNs were collected 6 hours or 1 day later for analysis. Approximately half of all B cells and T cells were found to be labeled by the Ab conjugate in this time course (fig. S4, F and G). A portion of these cells were stained with an anti-Cy5 antibody, and the ratio of anti-Cy5 (surface-bound eOD) to Cy5 signals (total eOD) revealed that most antigen was extracellular on both T and B cells at both time points (fig. S4, H and I). By staining another portion of the cells with VRC01 antibody and quantifying the ratio of VRC01 (intact CD4-binding site) to Cy5 signal (total eOD) compared with cells freshly labeled in vitro showed that ~80% of B cells carried eOD with an intact CD4-binding site at 6 hours, and this remained unchanged at 24 hours (Fig. 1, H and I). By contrast, eOD bound to T cells already trended toward a lower level of intact CD4 binding site at 6 hours, and T cell–bound eOD was 50% degraded by 1 day (Fig. 1, H and I). Thus, both FRET and antibody staining analyses suggest that antigens are rapidly degraded as they enter the SCS, interfollicular regions, and paracortex, but antigen localized to follicles, whether bound to FDCs or not, is protected from rapid proteolysis.

Metalloproteinases are expressed by sinus-lining and stromal cells and contribute to rapid extrafollicular antigen degradation

High-magnification imaging of LNs stained to identify cell membranes 24 hours after immunization revealed pockets of antigen localized in the interstices between cells, and FRET imaging showed that this antigen was degrading, suggesting a role for extracellular proteolysis (Fig. 2A). To quantify intracellular versus extracellular antigen breakdown in whole LNs, we developed an antigen pull-down assay, in which FRET dye–labeled antigens in the supernatants of mechanically disrupted LNs or lysates of isolated LN cells were captured on anti-Cy3/Cy5–conjugated beads, followed by flow cytometry–based FRET measurements (Fig. 2B). For this analysis, cells were excited at the donor dye wavelength and acceptor dye emission was collected, and this FRET signal was normalized by the total fluorescence obtained from exciting the acceptor and collecting its emission (fig. S5A). Total antigen recovered was quantified in parallel by measuring Cy5 fluorescence, calibrated by beads loaded with known quantities of Cy5-labeled eOD-60mer (fig. S5, B and C). These analyses revealed that the total amount of recoverable antigen dropped by 31-fold over 2 days, and at days 1 and 2, nearly equal proportions of antigen were extracellular versus intracellular (Fig. 2C). Focusing on the 6-hour and 1-day time points, when sufficient antigen was recovered for robust analysis, FRET signals showed that extracellular and intracellular antigen both underwent substantial degradation in this time frame (Fig. 2D and fig. S5D). Thus, extracellular and intracellular degradation contribute approximately equally to antigen clearance.

We next sought to determine factors governing extracellular antigen degradation. As shown by FRET analysis, eOD-60mer$_{40}$ incubated with mouse lymph or plasma in vitro remained intact (Fig. 2E). Thus, we examined the extracellular or secreted proteases expressed by LN cells by reanalyzing a single-cell RNA sequencing (scRNA-seq) dataset recently collected from LNs of naïve or immunized mice. We found >30 genes encoding extracellular or secreted proteases expressed in one or more LN cell types, predominantly in stromal cells and myeloid cells (Fig. 2E). Because we observed substantial antigen degradation at the SCS, we focused on sinus-lining SCS macrophages and lymphatic endothelial cells (LECs) (Fig. 2F). The most highly expressed genes in SCS macrophages (in both naïve and vaccinated mice) included *Mmp14*, *Adam17*, and *Adam10*, which encode metalloproteinases. We focused our attention on these three proteases because they were also expressed by LECs and were generally the most broadly and/or highly expressed across multiple cell types in the LN. *Mmp9* was also notable as the most highly expressed protease in LECs. Our dataset lacked FDCs, but we analyzed two published LN stroma scRNAseq datasets that included these cells (*28*, *29*). Consistent with our observations of limited antigen degradation in follicles, FDCs expressed very low to minimal levels of extracellular proteases compared with LN fibroblasts or LECs (fig. S5, E and F).

Immunohistochemistry analysis of ADAM17, ADAM10, MMP14, and MMP9 protein expression revealed high levels of these proteases along the subcapsular and medullary sinuses and within isolated regions in the interior of both resting and immunized LNs (Fig. 3A and fig. S6, A and B). By contrast, the FDC network and its immediate vicinity showed low levels of all four proteins. Consistent with the scRNAseq data, all three proteases partially colocalized with CD169$^+$ and LYVE1$^+$ cells, as shown by immunohistochemistry (Fig. 3B and fig. S6, C and D). We also examined the expression patterns of these proteases in human lymphoid tissues. Human tonsil tissue sections showed prominent ridges of ADAM17, MMP9, and MMP14 expression, with more scattered expression of ADAM10 (fig. S6E). However, CD35$^+$ follicles were largely devoid of all four proteases.

In vitro, the top three sinus-expressed proteases, ADAM17, ADAM10, and MMP14, all degraded eOD-60mer$_{40}$, although the FRET profile somewhat differed, likely due to unique patterns of cleavage for each metalloprotease (fig. S7A). We thus examined the relative localization of these proteases and degrading antigen in the sinuses. Six hours after injection, eOD-60mer$_{40}$ was predominantly colocalized with CD169$^+$ macrophages, followed by LYVE1$^+$ LECs and other cell types not identified by these two markers (fig. S7, B and C). The 60mer also showed substantial colocalization with ADAM17, ADAM10, and MMP14 (fig. S7D). Using high-magnification FRET imaging on immunostained LN sections, we found that ~60% of degraded eOD-60mer$_{40}$ colocalized with at least one of these three metalloproteinases at 24 hours after immunization (Fig. 3C).

To functionally test the role of extracellular protease activity in antigen breakdown, we prepared live vibratome sections of LNs isolated from mice 2 hours after immunization with eOD-60mer$_{40}$ (a time point when most antigen in the LN is still intact) and incubated tissue slices in the presence or absence of protease inhibitors. In LN slices cultured without inhibitors, eOD-60mer degradation proceeded, with ~60% of antigen intact relative to the starting material after 6 hours, whereas LNs cultured with the metalloproteinase-specific inhibitor marimastat reduced this degradation by ~50% (Fig. 3D). Moreover, media containing marimastat supplemented with a broad-spectrum protease inhibitor cocktail did not yield significantly more intact antigen (Fig. 3D). We next treated mice with marimastat before and during immunization with eOD-60mer$_{40}$. As shown in Fig. 3E, marimastat treatment led to a 40% increase in the total amount of intact extracellular antigen recovered by our antigen pull-down assay 24 hours after immunization compared with vehicle control. In parallel, microscopy-based FRET analysis of LN sections similarly showed that significantly more eOD-60mer$_{40}$ was intact within the SCS after marimastat treatment (Fig. 3E). Collectively, these data suggest that proteases, particularly metalloproteinases, contribute to rapid extrafollicular antigen breakdown in LNs.

Extracellular protease activity is enriched in sinus-lining macrophages and stromal cells but low within follicles

We next sought to understand why antigen localized to FDCs was protected from proteolysis. We observed low staining of ADAM17, ADAM10, and MMP14 in follicles, but this data could not rule out expression of other proteases in the vicinity of FDCs. We thus used an imaging zymography approach to visualize protease activity more comprehensively on live tissue sections (Fig. 4A) (*30*). Live LN vibratome sections were incubated with two fluorescent peptide probes. The first, an

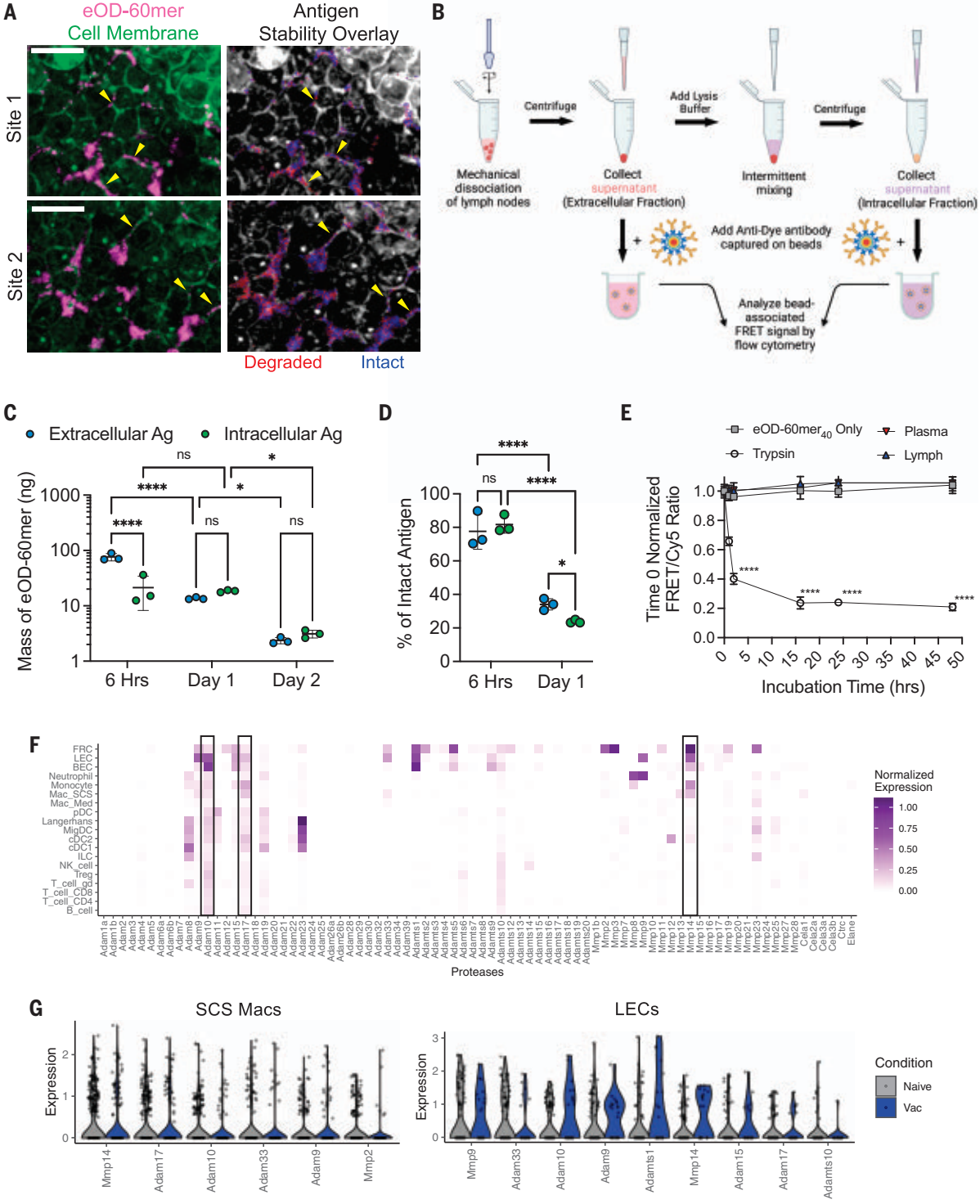


Fig. 2. After immunization, a significant proportion of antigen remains extracellular and is degraded over time, correlating with the expression of ADAM and MMP proteases by macrophages and stromal cells in the LN. (**A**) C57BL/6 mice were immunized with 5 μg of saponin adjuvant and 10 μg of eOD-60mer$_{40}$ and harvested 1 day later for sectioning and confocal imaging. Two representative extrafollicular regions are shown stained with the cell membrane dye CellMask Green (left panels) and an overlay of false color binary FRET images indicating intact (blue) and degraded (red) eOD-60mer$_{40}$ (right panels). Scale bars, 10 μm. (**B** to **D**) C57BL/6 mice were immunized as in (A) and draining LNs were isolated (n = 3 pools/time point, with each pool containing four LNs from two mice) for measurement of extracellular versus intracellular antigen quantity and integrity by flow cytometry–based FRET analysis. Shown are a schematic illustrating the workflow for antigen pull-down assay (B), total mass of antigen recovered from whole LN (C), and percentage of intact antigen (D) retrieved at the indicated time points after immunization. ns, not significant; *q ≤ 0.05; ****q ≤ 0.0001 by two-way ANOVA for (C) and (D) with two stage step-up method to correct for false discovery rate during multiple comparisons. (**E**) FRET analysis of eOD-60mer$_{40}$ integrity after incubation with 10% diluted plasma or lymph from naïve C57BL/6 mice or trypsin at 37°C (n = 3 replicates/group). P values were determined by one-way

ANOVA followed by Tukey's post test. (**F** and **G**) scRNAseq analysis of protease expression in LNs from naive C57BL/6 mice or mice immunized 6 hours earlier with ovalbumin peptide and CpG. (F) Heatmap of average normalized expression of extracellular protease genes in single cells across LN cell types in immunized mice. (G) Normalized expression levels of the most highly expressed proteases in SCS macrophages and LECs arranged in descending order of average expression; protease genes with an average expression value >0.1 are included. All plots show mean ± SD.

activatable zymography probe (AZP), is composed of a 5-carboxyfluorescein dye linked to a cationic polyarginine (polyR) peptide, followed by a broad-spectrum protease-cleavable substrate sequence and an anionic polyglutamic acid (polyE) peptide that complexes with the polyR sequence. When proteases cleave the substrate site of the AZP, the polyR sequence is freed and electrostatically binds to nearby cells, labeling the site of protease activity (Fig. 4A). The AZP selected was cleaved by metalloproteinases including ADAM17, ADAM10, and MMP14, as well as aspartic and cysteine protease cathepsins reported to be extracellularly active and expressed by immune cells (fig. S7E) (*30*–*32*). A second control probe, a Cy5-conjugated polyR peptide, provides a map of overall electrostatic-binding sites in the tissue. Tissue areas labeled with both probes identify proteolytically active sites, whereas locations labeled by the control peptide but not the AZP indicate areas devoid of protease activity.

Confocal imaging of LN tissue sections from naïve or immunized animals showed AZP labeling of discrete cells across the entire tissue, with particular concentrations of protease activity at the SCS, except within the immediate vicinity of the FDC networks (Fig. 4B and fig. S7F). Incubation of LN sections with a control uncleavable D-isomer form of the same AZP peptide sequence showed substantially lower fluorescence signal, suggesting that AZPs were not internalized and required extracellular proteolysis to bind to cells (fig. S7, G to I). Quantification of AZP signal within FDC regions versus adjacent LN tissue revealed substantially lower protease activity in the follicles of both resting and immunized LNs (Fig. 4C).

Flow cytometry analysis of single cells extracted from polyR- and AZP-treated LN sections revealed that 80% or more of all lymphocyte and stromal cell populations examined were positive for control polyR binding (fig. S8, A to C) (*33*). Gating on cells positive for polyR binding (indicating sufficient binding sites for AZP-based protease detection), we found that neutrophils, SCS macrophages, and fibroblastic reticular cells had the highest levels of AZP binding above background labeling by D-isomer control AZP in both immunized (Fig. 4, D and E, and fig. S8D) and resting LNs (fig. S8E). All macrophage subsets examined ($CD169^+CD11b^{hi}F4/80^{low}$ subcapsular; $CD169^+CD11b^{low}F4/80^{low}$ interfollicular; and $CD169^+F4/80^{hi}$ medullary macrophages) had some proportion of AZP^+ cells, but SCS macrophages showed the most prominent protease activity (Fig. 4, D and E, and fig. S8D). Lymphatic endothelial cells (LECs), DCs, and monocytes also exhibited lower but clearly detectable protease activity, whereas FDCs, B cells, and T cells showed very low levels of AZP labeling (Fig. 4, D and E).

We next assessed other potential mechanisms of antigen protection within follicles. Because eOD-60mer is recognized by MBL and becomes complement decorated in the presence of serum, it seemed possible that MBL and/or complement binding might have a shielding effect by blocking proteolytic attack on FDC-localized antigen. However, in an in vitro trypsin digestion assay, eOD-60mer was proteolyzed with destruction of the VRC01 epitope regardless of serum-derived MBL or C3 binding on the NP surface (fig. S9A). A second possibility was that antibodies produced very early after immunization form immune complexes to sterically inhibit protease attack, or that FDCs capturing immune complexes rapidly internalize and recycle it on their dendrites, providing a degree of protection from protease exposure in the follicles. We immunized MD4 mice in which B cells transgenically express an antigen receptor specific for an irrelevant antigen with $eOD\text{-}60mer_{40}$, and found that antigen localized to FDCs after 2 days was largely nondegraded, which was similar to what we found in wild-type mice, suggesting that antigen protection is not mediated by antibodies at least during these early time points (fig. S9B).

To assess whether FDCs could protect antigen in the face of artificially enforced protease activity in follicles, we immunized mice with $eOD\text{-}60mer_{40}$ and adjuvant, extracted LNs 3 days later, when antigen was concentrated on the FDC networks, and incubated live LN slices with trypsin or recombinant matrix metalloproteinase 9 (MMP9) to expose FDC-localized antigen to protease attack. Both trypsin and MMP9 treatment led to a significant loss of intact eOD-60mer within 4 hours, suggesting that antigen localized to FDCs is not intrinsically protease resistant (Fig. 4, F and G). To determine whether low protease activity in follicles was important for the retention of intact antigen in vivo, we devised a strategy to introduce active extracellular proteolysis adjacent to FDCs. For this, polyclonal B cells were transduced to express a constitutively active form of MMP9 (*34*), together with mNeonGreen as a reporter or mCherry reporter only as a control (fig. S10, A to D), and injected intranodally into mice that had been immunized 2 days earlier with $eOD\text{-}60mer_{40}$ (Fig. 4H). Two days after B cell transfer, LNs were recovered for FRET imaging. As shown in Fig. 4I, engineered B cells dispersed into follicles in and around the FDC networks. Image analysis revealed that transfer of MMP9-expressing B cells led to 65% less antigen retained on FDCs after 2 days (Fig. 4, J and K), and that 37% less of the retained antigen was intact (Fig. 4L). Therefore, although additional factors may play a role, the lack of active protease activity in the vicinity of FDCs appears to be an important contributor to the long lifetime of immunogens captured in the FDC networks.

Antigen localization to FDCs selectively enhances B cell responses to intact antigen but not breakdown products

We hypothesized that rapid antigen degradation in extrafollicular regions of the LN could limit responses to the intact antigen, whereas rapid delivery of antigen to FDCs might enhance responses to the native immunogen. To test this idea, we performed experiments using a stabilized HIV Env gp140 SOSIP trimer called N332-GT2 (*35*) in soluble trimer and protein nanoparticle forms. Labeling conditions were identified (six dyes per trimer on average, $trimer_6$) that allowed for FRET tracking of N332-GT2 integrity with minimal perturbation of the antigenicity profile of the trimer (fig. S11A). Similar to our findings with eOD-60mer, the *E* of protease-treated $trimer_6$ was highly correlated with the proportion of intact antigen, as determined by PAGE analysis in vitro (fig. S11, B to D). Loss of FRET signal also correlated with reduced binding of structure-sensitive antibodies (fig. S9E). Concomitantly, antibodies that recognize epitopes of the gp120 inner domain buried in the interior of properly folded trimers bound to trypsin-digested $trimer_6$ (fig. S9E). The trimer was stable in serum or lymph for at least 48 hours, although it was degraded by trypsin or recombinant MMP9 (fig. S9F). This trimer was also produced as an NP by fusion of the trimer sequence with ferritin subunits (abbreviated trimer-NPs) (*35*). For N332-GT2 nanoparticles, 40 dyes could be conjugated to each particle ($trimer\text{-}NP_{40}$) without affecting the binding of trimer structure–specific mAbs (fig. S9G).

We first examined the localization and in vivo stability of trimers and trimer-NPs after immunization (Fig. 5A). We previously showed that ferritin-based trimer-NPs are rapidly transported to FDCs through the same MBL/complement pathway as the eOD-60mer immunogen, whereas soluble trimer distributes more diffusely in

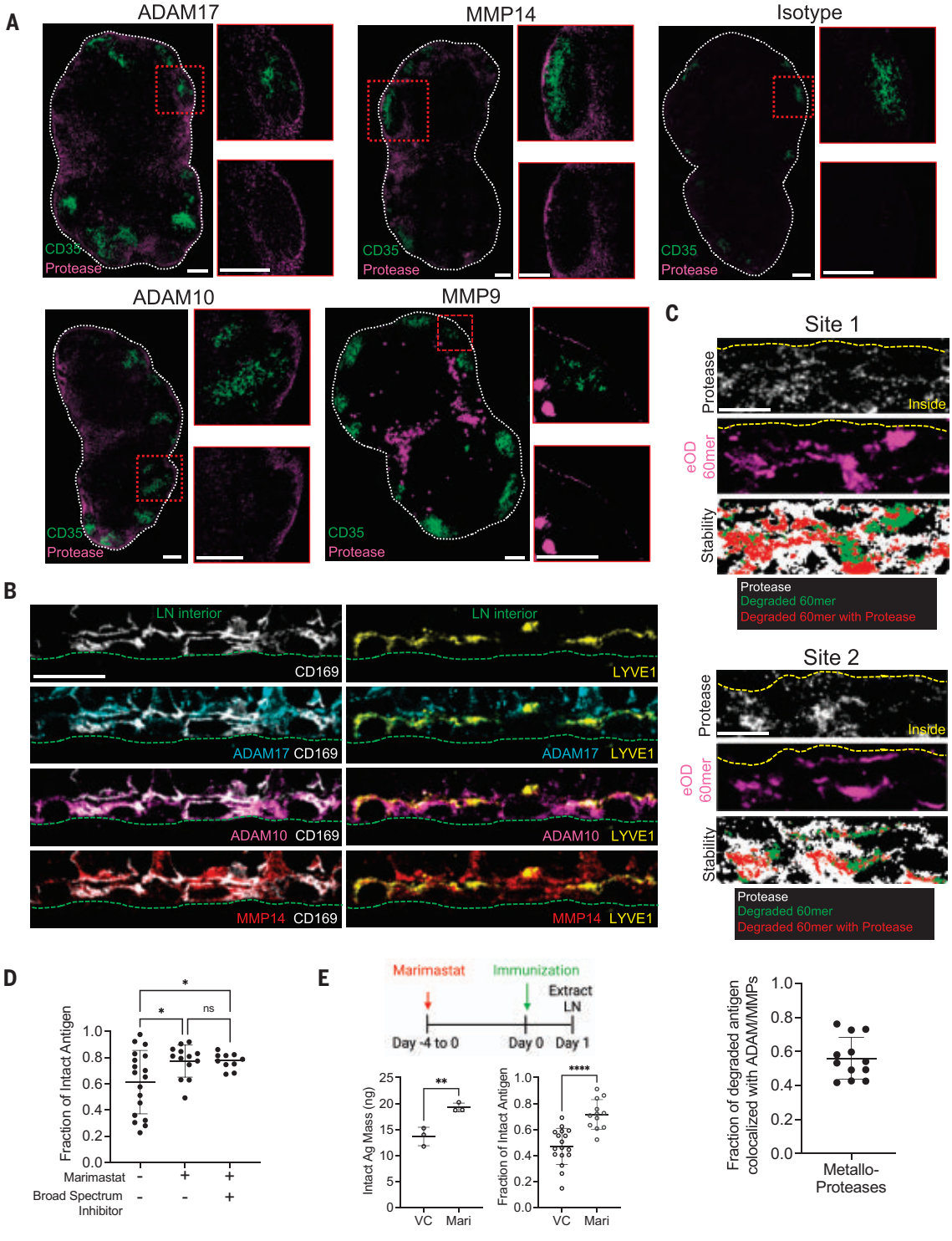


Fig. 3. Metalloproteinases are expressed by sinus-lining cells and contribute to rapid antigen degradation in LNs. (**A** and **B**) Resting LNs from C57BL/6 mice (n = 3 animals/group) were flash frozen and cryosectioned. (A) Anti-CD35 (green) and anti-metalloproteinase or isotype control mAbs (magenta) staining of LN sections and magnified views of the red dashed regions of interest. Scale bar, 200 μm. (B) Magnified images of the SCS stained for CD169$^+$ macrophages (left column) or LYVE1$^+$ LECs (right column) and the indicated proteases. Scale bar, 20 μm. (**C**) C57BL/6 mice (n = 3 animals) were immunized with 5 μg of saponin adjuvant and 10 μg of eOD-60mer$_{40}$. After 24 hours, LNs were harvested, sectioned, and collectively stained for ADAM10, ADAM17, and MMP14 before FRET imaging. Two representative SCS regions showing protease expression, eOD-60mer localization, and false color overlay of proteases (white), total degraded antigen (green), and degraded antigen colocalized with proteases (red) are shown. Scale bar, 20 μm. Bottom graph quantifies degraded eOD-60mer colocalized with metalloproteinases. Each point represents one region, and data were collected from at least six tissue sections from six LNs. LN boundaries are indicated by white (A), green (B), or yellow (C) dashed lines. (**D**) C57BL/6 mice (n = 3 animals/group) were immunized as in (C) and LNs were collected after 2 hours before vibratome slicing. Live LN slices were left untreated or incubated in marimastat (Mari) with or without broad-spectrum protease inhibitors for 6 hours ex vivo before FRET imaging. Each point represents one region, and data were collected from at least 10 tissue sections from six LNs. P values were determined by one-way ANOVA with Tukey's post test. (**E**) C57BL/6 mice were treated with Mari for 5 days and then immunized as in (C), and LNs were analyzed 24 hours afterward by pull-down assay (n = 3 pools/group, each pool containing four LNs from two mice) or FRET imaging (n = 3 animals/group). Shown are Mari or vehicle control (VC) dosing and immunization schedule (top panel), intact extracellular antigen recovered by pull-down assay (lower left panel), and FRET imaging analysis for intact antigen (bottom right panel). Each point represents one region for FRET imaging, and data were collected from at least six tissue sections from six LNs. P values were determined by Student's t test. All plots show mean ± SD.

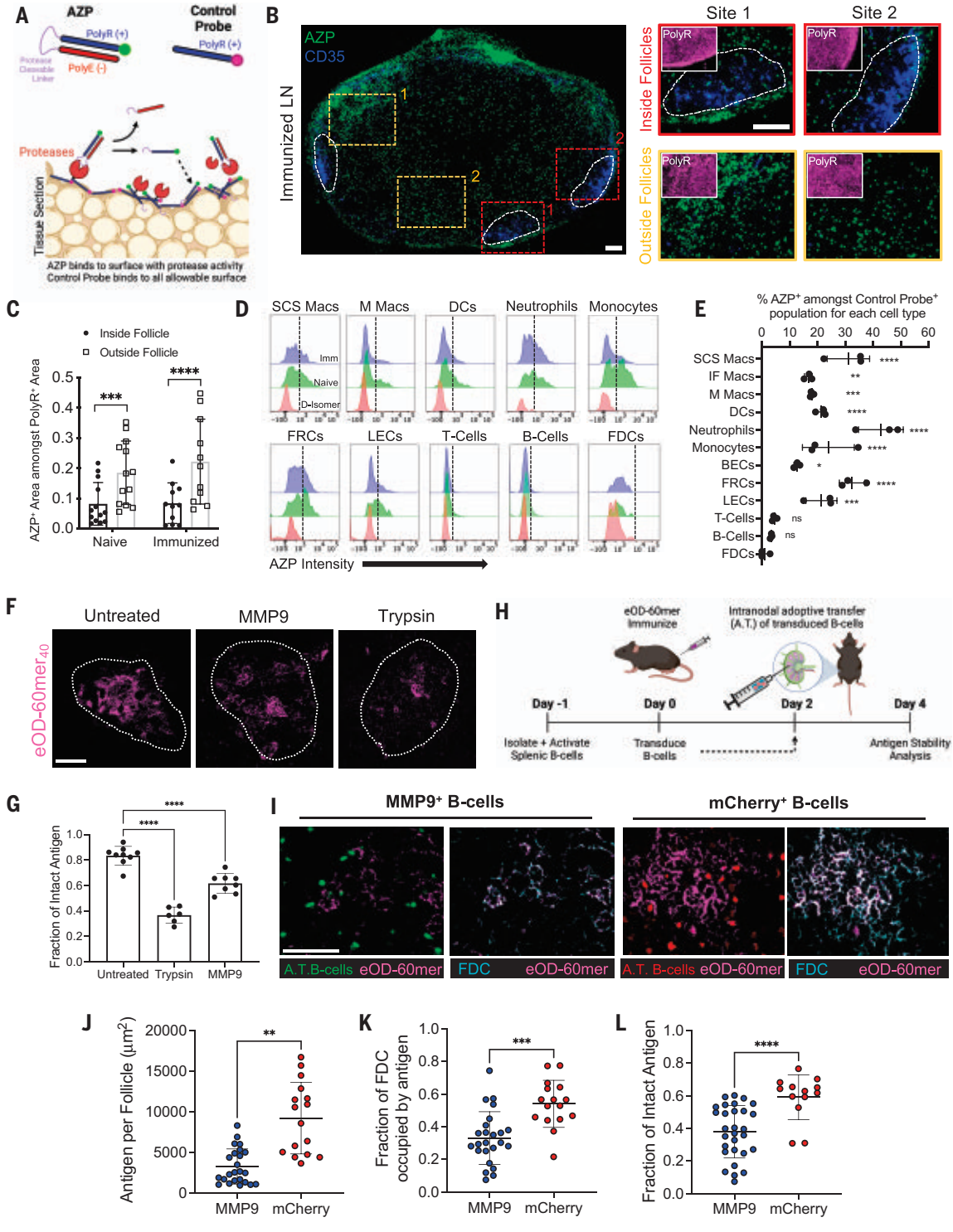


Fig. 4. Protease activity is spatially heterogeneous in LNs, with high levels in the SCS and low activity in B cell follicles. (**A**) Schematic illustrating AZPs and protease activity. (**B** to **E**) Live LN slices from C57BL/6 mice (n = 3 animals/group), either naïve or immunized with eOD-60mer and 5 μg of saponin adjuvant, were incubated with AZP and control probes. (B) Confocal microscopy of LN slices (left) showing AZP$^+$ cells (green), CD35$^+$ FDCs (blue), and magnified views (right) of the boxed regions "Inside Follicles" or "Outside Follicles." Insets show control polyR probe binding (magenta). White contours denote follicles. (C) Quantification of AZP binding inside follicles versus outside follicles. Each point represents one LN section collected from a total of 11 to 13 sections from six LNs. P values were determined by Student's t test. (D) Representative flow histograms of AZP and D-isomer probe staining for cells from immunized (Imm) or resting (Naïve) LNs. Dashed lines denote AZP$^+$ and AZP$^-$ populations. (E) Quantification of AZP$^+$ cells. P values were determined by one-way ANOVA with Dunnett's post test against FDCs. (**F** and **G**) C57BL/6 mice (n = 3 animals/group) were immunized with 5 μg of saponin adjuvant and 10 μg of eOD-60mer$_{40}$. Live LN slices were generated 3 days later and incubated with trypsin, MMP9, or left untreated. Representative images of eOD-60mer$_{40}$ remaining on FDCs (F) and fraction of intact antigen on FDCs based on FRET imaging (G). White lines show follicle boundaries. Scale bar, 100 μm. Each point represents one region from one tissue section. Data were collected from at least six tissue sections from six LNs. P values were determined by one-way ANOVA with Dunnett's post test against untreated sections. (**H** to **L**) C57BL/6 mice (n = 3/group) were immunized as in (F) and 2 days later, MMP9- or mCherry-expressing polyclonal B cells were adoptively transferred. (H) Timeline for retroviral transduction, adoptive transfer, and analysis. (I) Images of transferred MMP9$^+$ (left panel) and mCherry$^+$ control B cells (right panel) near FDCs bearing eOD-60mer$_{40}$. Scale bar, 100 μm. (J to L) Image analysis for total eOD-60mer$_{40}$ in the follicle (J), fraction of FDC network with eOD-60mer$_{40}$ (K), and integrity of FDC-localized eOD-60mer$_{40}$ (L). P values were determined by Student's t test. All plots show mean ± SD.

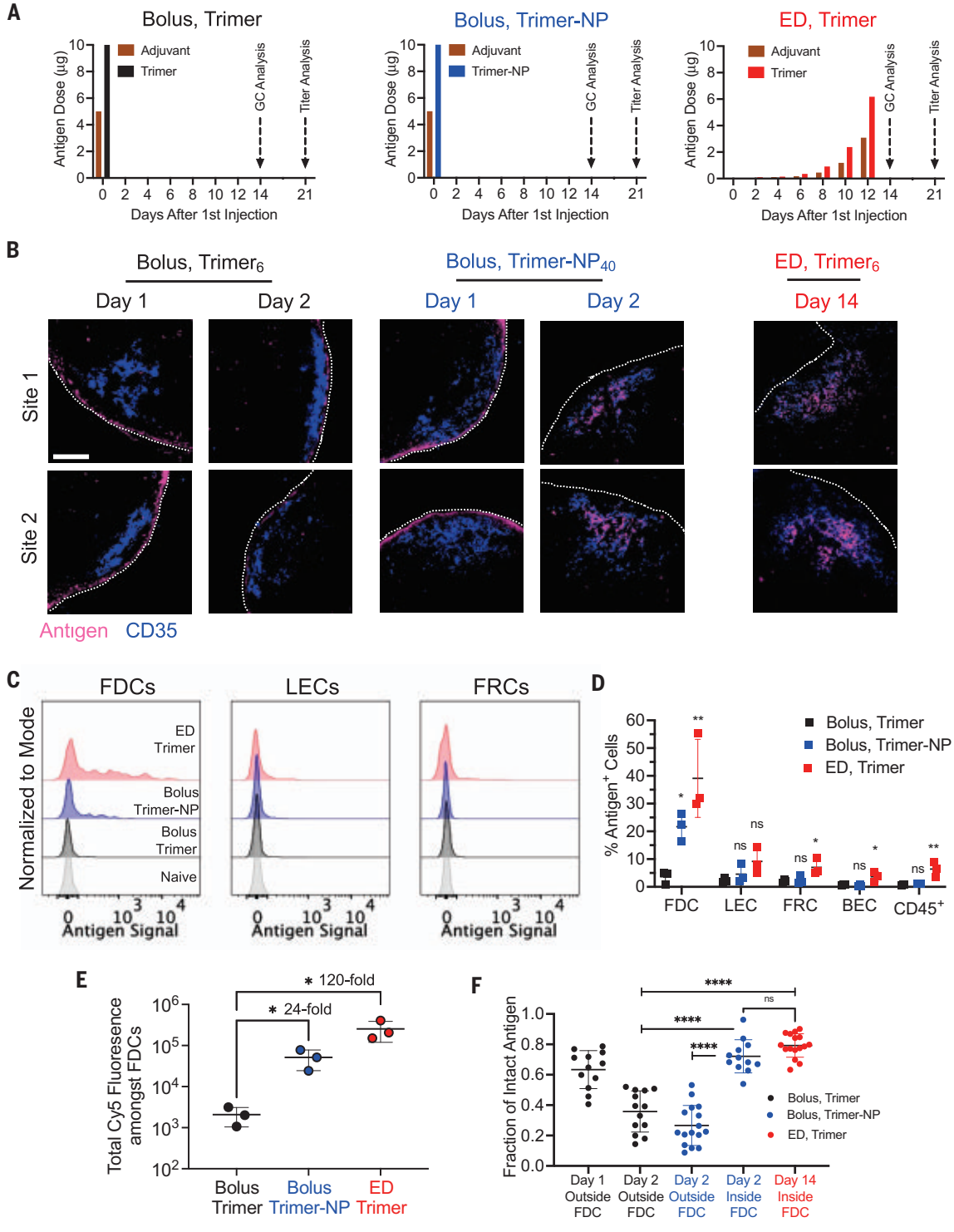


Fig. 5. Immunizations targeting antigen to FDCs preserve antigen integrity in vivo. (**A**) Administration schedules for bolus and escalating dose (ED) immunization regimens using soluble trimer and trimer-NP immunogens. Saponin adjuvant was used for all conditions. (**B**) C57BL/6 mice (n = 3 animals/group) were immunized with indicated dye-labeled antigens and dosing scheme. At the indicated time points, LNs were flash frozen, cryosectioned, and imaged for antigen localization. Shown are representative images of two distinct follicles illustrating antigen (magenta) and CD35 (blue). Scale bar, 100 μm. (**C** to **E**) Flow cytometry analysis of LN cells (n = 3 pools/group, with each pool containing four LNs from two mice) isolated after immunization with fluorescently labeled trimer (day 2 for bolus trimer and trimer-NP injection, or day 14 after the initial injection of ED trimer). Shown are representative histograms of antigen intensities among LN cells (C), quantification of the antigen$^+$ population (D), and the sum of all antigen intensities among all antigen$^+$ FDCs (E) for the indicated immunization conditions. P values were determined by one-way ANOVA with Dunnett's post test for pairwise comparisons against bolus trimer. (**F**) C57BL/6 mice (n = 4 animals/group) were immunized as in (B) and LNs were prepared for FRET imaging at the specified time points. The fractions of intact antigens in the LN at the indicated locations and times are shown. Each point represents one region from one tissue section. Data were collected from at least eight tissue sections from eight LNs. P values were determined by one-way ANOVA with Tukey's post test. All plots show mean ± SD.

LNs before clearance (*22*, *36*). Here, bolus-injected soluble trimer$_6$ or trimer-NP$_{40}$ were detected in the SCS for up to 2 days, but trimer-NP$_{40}$ also accumulated on FDCs starting 2 days after immunization (Fig. 5B). We also tested an "escalating-dose" (ED) immunization (*10*, *37*), giving the same total trimer$_6$/adjuvant dose but administered as seven injections of increasing dose over 2 weeks (Fig. 5A). ED immunization is another approach that promotes antigen deposition on FDCs through immune complex formation as the dosing progresses (*10*, *37*). Two days after the last ED injection (day 14) of trimer$_6$, substantial antigen was localized on the FDC network (Fig. 5B).

Flow cytometry analysis of LNs pooled from multiple mice revealed that FDCs were by far the major cell type where antigen accumulated after immunization (Fig. 5, C and D, and fig. S12A), and trimer-NP or ED immunizations increased the amount of FDC-trapped antigen by 24-fold and 120-fold over bolus immunization, respectively (Fig. 5E). FRET analysis indicated that substantial portions of extrafollicular antigen (60% of trimer and 80% of trimer-NP) were degraded within 2 days, whereas trimer-NP_{40} or ED-administered $trimer_6$ localized to FDCs remained predominately intact (Fig. 5F). Thus, immunization with HIV Env trimers using ED or nanoparticle regimens promotes follicular targeting and increases the level of intact antigen retained within the LN.

Bolus immunization with trimer-NP or ED immunization with trimer led to increased total numbers of GC B cells compared with bolus trimer immunization, by 8- and 27-fold, respectively (Fig. 6, A and B). Staining with N332-GT2 trimer probes to identify cells capable of binding to the intact antigen at day 14 revealed ~200-fold and ~40-fold increases in the number of GC B cells specific for the intact trimer after ED immunization and trimer-NP immunization, respectively, compared with bolus trimer injection (Fig. 6B). Trimer probe staining of GC B cells from mice immunized with an irrelevant antigen, ovalbumin, showed low background staining (Fig. 6B). The proportion of trimer-specific GC B cells after bolus trimer vaccination remained relatively constant through 21 days after immunization and never reached the levels primed by trimer-NP or ED vaccination (fig. S12B).

We also investigated the frequency of GC B cells reacting to trimer breakdown products by staining GC B cells with HR1, C1, or V3 loop peptide probes, representing epitopes buried in the intact trimer (Fig. 6, C and D, and fig. S12C). Fourteen days after immunization, low but statistically significant levels of GC B cells recognizing the C1 and V3 epitopes were detectable after bolus trimer immunization (Fig. 6D and fig. S12C), and collectively more B cells targeted these three breakdown product epitopes than intact trimer (Fig. 6E). Trimer-NP and ED immunizations greatly amplified the on-target GC B cell response against intact trimer, but had minimal effect on the magnitude of GC responses targeting breakdown product epitopes (Fig. 6D and fig. S12C). Further, compared with bolus trimer immunizations, NP or ED yielded 1000-fold higher serum immunoglobulin (IgG) titers recognizing the intact trimer 3 weeks after immunization (Fig. 6F and fig. S12D). Low titers against intact trimer observed after bolus trimer immunization were of the same order of magnitude as responses to the HR1/C1/V3 breakdown product epitopes (Fig. 6F and fig. S12D).

To determine whether having a low level of protease activity in follicles is important for the efficacy of FDC-targeting immunizations, we again adoptively transferred MMP9-expressing polyclonal B cells to introduce protease activity into the follicles. MMP9-expressing or control mCherry-expressing B cells were injected intranodally 2 days after immunization with trimer-NPs, and 12 days later, GC responses were analyzed (Fig. 6G). As shown in Fig. 6, H and I, introduction of $MMP9^+$ B cells did not affect the total GC response, but GC B cells binding to the intact trimer were reduced by sevenfold, and HR1 peptide-binding cells were increased by about twofold. Thus, low protease activity in follicles is important for optimal on-target B cell responses to FDC-targeted antigens.

Enhancements in GC responses targeting the intact immunogen elicited by NP or ED immunization likely reflect, in part, the fact that GC B cells are encountering antigen in a multivalent form (either the NP or as immune complexes in the case of ED immunization) and/or costimulatory effects of complement decorating the antigen with these immunizations. To test the impact of FDC localization in the absence of these additional factors, we performed immunizations targeting monomeric eOD antigen to FDCs by creating a fusion of eOD and an anti-CD35 ScFv (αCD35-eOD) or an isotype control ScFv (Iso-eOD) (fig. S12E). αCD35-eOD accumulated on FDCs within 2 days after immunization, whereas isotype-eOD remained localized primarily in the sinus and extrafollicular regions (Fig. 6, J and K). Immunization with αCD35-eOD elicited about fourfold more total GC B cells and about ninefold more eOD-specific GC B cells than Iso-eOD (Fig. 6, L and M, and fig. S12, F and G). Moreover, serum IgG titers at 21 days after injection were 11-fold higher for the FDC-targeted antigen (Fig. 6K and fig. S12G). Therefore, FDC localization can have a substantial impact on the priming of B cells against intact antigen structures even for monomeric antigens.

Discussion

Elicitation of protective antibody responses by vaccination is predicated on the activation of B cells with antigen receptors that recognize protective epitopes present on a targeted pathogen. Elegant advances in protein engineering have led to the generation of vaccine immunogens that faithfully mimic native viral envelope proteins and exhibit greatly enhanced structural stability relative to the native antigens (*11*, *38*). However, despite the critical importance of immunogen structure to the specificity of the resulting immune response, the lifetime of structurally intact antigen in vivo after immunization is poorly understood. Here, we applied a FRET imaging–based approach to track the fate of protein immunogens in mice, and made several important findings. First, we found that LN protease activity is spatially heterogeneous, with high protease levels and activity found in stromal cells and macrophages, especially in the sinuses, but low levels of protease activity in B cell follicles. Second, consistent with this observation, antigens arriving in the SCS and extrafollicular regions of LNs rapidly begin degrading, with significant losses in integrity within 24 to 48 hours. By contrast, immunogens localized to FDCs within this early time window remain largely intact and are preserved for more than a week. As a result of this spatial compartmentalization of protease activity, traditional bolus immunization using soluble HIV envelope antigens, which does not efficiently deliver antigen to B cell follicles, led to weak GC B cell responses to the structurally intact immunogen, which were matched in magnitude by responses against irrelevant breakdown products. However, using immunizations that promoted antigen accumulation on FDCs, responses to intact antigen could be greatly amplified without concomitant increases in responses to off-target antigen breakdown products.

We hypothesize that spatially heterogeneous protease activity within LNs is part of an evolutionary strategy to halt pathogen dissemination at draining LNs before systemic spread while preserving the ability to produce effective structure-specific antibodies if needed. In this concept, LNs serve a first line of defense function as immediate pathogen filtration and elimination units, with adaptive immunity developing over time in parallel only if immediate microbe clearance cannot be achieved. This idea was first proposed in the setting of local viral or bacterial infections, in which draining LN innate immune responses were important for blocking systemic dissemination of pathogens (*39*). Extracellular LN protease activity may similarly contribute to attenuating infection, because it is known that metalloproteases can inactivate some viruses (*40*, *41*). However, the adaptive immune response also needs to be capable of "seeing" functional pathogen epitopes to generate useful antibody responses, motivating the necessity for a site where B cells can encounter intact antigen.

Although protease activity in LNs is understudied, antigen degradation within the SCS has been reported. Jenkins and colleagues observed that hen egg lysozyme conjugated to polymer microspheres could be cleaved and acquired by hen egg lysozyme–specific B cells within 4 hours of injection, which the authors ascribed to proteolytic activity in lymph (*12*). However, we found that incubation of eOD-60mer or stabilized HIV Env trimers in plasma or lymph resulted in minimal degradation of the antigens. Healthy peripheral lymph has been shown by proteomic analysis to contain a relatively narrow set of proteases, including collagenase type IV, carboxypeptidase B2, and

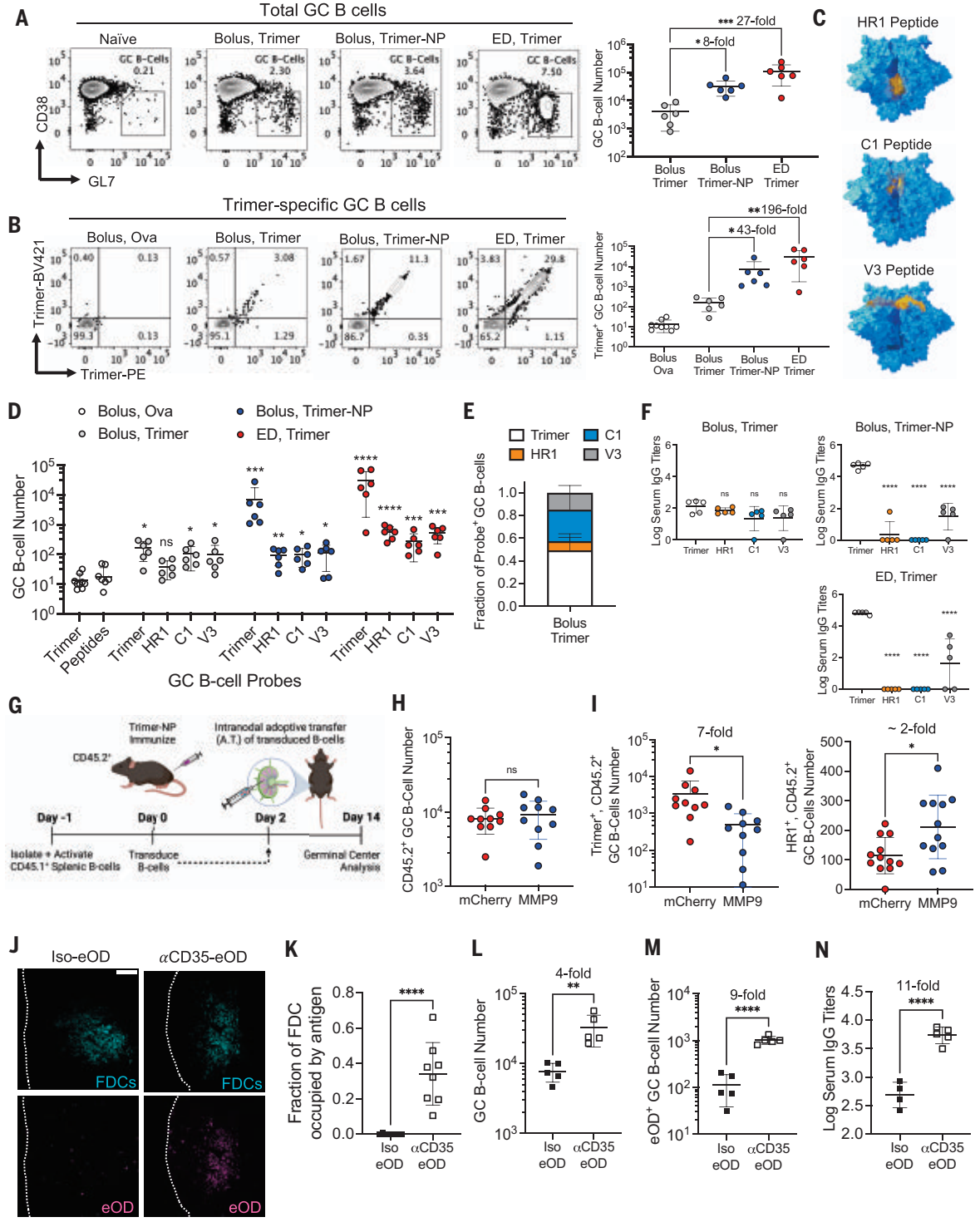


Fig. 6. Targeting of immunogens to FDCs amplifies GC B cell responses against intact antigens without increasing the response to antigen breakdown products. C57BL/6 mice (n = 5 to 10 animals/group) were immunized with the indicated antigens and saponin adjuvant according to the dosing schemes. (**A** to **E**) LNs were harvested at day 14 for flow cytometry analysis. (A) Representative flow plots showing GC B cells (left) and their quantification (right). (B) Representative flow plots of intact trimer-specific GC B cells (left) and their quantification (right). *$q \le 0.05$; **$q \le 0.01$; ***$q \le 0.001$ by Kruskal-Wallis one-way ANOVA followed by two stage step-up method for comparison against bolus trimer. (C) Structural models of the HIV Env trimer highlighting (in yellow) the location of the HR1, C1, and V3 peptides used as probes to identify breakdown product–specific GC B cells. (D) Enumeration of intact trimer and off-target peptide specific GC B cells after bolus, trimer-NP, ED trimer immunization, or bolus ovalbumin-immunized mice (control). ns, not significant; *$q \le 0.05$; **$q \le 0.01$; ***$q \le 0.001$; ****$q \le 0.0001$ by Kruskal-Wallis one-way ANOVA with two stage step-up method for comparison against bolus ovalbumin. (E) Distribution of GC B cells recognizing indicated probes after bolus trimer immunization. (**F**) Serum IgG titers against intact trimer or breakdown product on day 21. P values were determined by one-way ANOVA and Dunnett's post test against intact trimer. (**G** to **I**) MMP$^+$ or control mCherry$^+$

polyclonal B cells were adoptively transferred intranodally into mice (n = 10 animals/group) immunized with 10 μg of trimer-NP and 5 μg of saponin adjuvant, followed by flow analysis. Shown are the experimental timeline (G) and enumeration of total host CD45.2$^+$ GC B cells (H), intact trimer [(I), left], and HR1 peptide specific GC B cells [(I), right]. P values were determined by Student's t test. (**J** to **M**) C57BL/6 mice (n = 4 to 5 animals/group) were immunized with 2.5 μg of Iso-eOD or αCD35-eOD and 5 μg of saponin adjuvant, followed by histology and flow cytometry. Shown are images of labeled antigen in the FDC networks (J) and its quantification (K) 2 days after immunization. Scale bar, 100 μm. (**L** to **N**) Flow cytometry analysis enumerating GC B cells (L) and eOD-specific GC B cells at day 14 (M), and eOD-specific serum IgG titers at day 21 (N). P values were determined by Student's t test. All plots show mean ± SD.

carboxypeptidase N (*42*, *43*), representing a much narrower range of proteases than we detected in the LN itself.

Antigen retention and protection within FDCs was first suggested more than three decades ago in seminal work by Tew *et al*. Immunogen stability was shown indirectly through chromatography after retrieval of radiolabeled antigen from whole LNs (*13*). However, with the emerging evidence that lymphatic endothelial cells can also store antigen in a complement receptor 2–independent manner, antigen retrieved from whole LN extracts may not be specific to FDCs (*44*). mAbs have been used to detect antigens within FDCs (*45*), but this approach can only report on individual epitopes and is limited by the potential for vaccine-elicited antibody responses to confound interpretation. Experiments such as these have suggested that the FDC network provides a sanctuary site that not only retains antigen but also protects it from degradation (*46*); however, direct evidence in support of this idea has remained sparse. Here, AZP-based spatial zymography assays applied to live tissue sections suggested that extracellular protease activity is low in follicles relative to the SCS and interfollicular regions of LNs. Prior work has shown that FDCs cyclically present and internalize captured antigens from their surfaces, which would be expected to aid in protecting antigen locally (*15*). However, when we challenged LN sections with recombinant proteases or transferred protease-expressing B cells into follicles in vivo, FDC-localized antigen underwent degradation. Nevertheless, it seems likely that antigen recycling and low protease activity in follicles work together to preserve antigen on FDCs over prolonged periods.

In summary, the studies presented here provide evidence for rapid proteolysis and clearance of antigens from extrafollicular regions of LNs after immunization, but also show long-lived retention of structurally intact antigen by follicular dendritic cells. In the absence of efficient follicle targeting, a significant proportion of B cells specific for irrelevant breakdown products are primed by vaccination. Such a substantial response against irrelevant epitopes may be problematic for difficult pathogens such as HIV, in which B cells capable of maturing to produce broadly neutralizing antibodies are very rare (*19*), because competition in GCs can overwhelm these rare precursors (*4*, *6*). These data provide clear motivation for immunization strategies that promote antigen delivery to the FDC network. Although we have demonstrated two approaches to achieve this, nanoparticle delivery and escalating dosing immunization, a variety of approaches may also be capable of achieving this goal. These include immunization with immune complexes (*47*), use of antigen-complement fusions (*48*), and extended antigen delivery through nucleic acid vaccines (*49*). Such approaches may be particularly important for the generation of protective antibody responses against difficult pathogens such as HIV and for the generation of broadly neutralizing antibody responses to other variable pathogens.

Materials and Methods

Mice

Six- to 9-week-old female C57BL/6 mice were purchased from The Jackson Laboratory (strain no. 000664). Tg(IghelMD4)4Ccg/J breeding pairs were purchased from The Jackson Laboratory (stock no. 002595) and mated in-house. Their progeny was genotyped (Transnetyx), and female mice between the ages of 8 and 11 weeks were used for immunization studies.

Saponin adjuvant preparation

The saponin adjuvant used in this study is an ISCOM-like self-assembled nanoparticle composed of cholesterol, phospholipid, and *Quillaja* saponin as described previously (*22*). Briefly, under sterile conditions, the following solutions were made: 0.5 ml of 20 mg/ml cholesterol (700000P, Avanti) in 20% MEGA-10 detergent (D6277, Sigma-Aldrich), 0.5 ml of dipalmitoylphosphatidylcholine (DPPC, 850355C, Avanti) in 20% MEGA-10 detergent, and 0.5 ml of 100 mg/ml Quil-A adjuvant (vac-quil, Invivogen) in deionized H_2O. DPPC solution was mixed with the cholesterol, followed by the addition of Quil-A saponin in rapid succession. This mixture was diluted with phosphate-buffered saline (PBS) to a concentration of 1 mg/ml cholesterol and 2% MEGA-10 before overnight equilibration at 25°C. The lipids/saponin/surfactant solution was then dialyzed against PBS using a 10-kDa molecular weight cut-off (MWCO) membrane for 5 days at 25°C and filter sterilized using a 0.2 Supor syringe filter. For further purification, the adjuvant solution was concentrated using 50-kDa MWCO Amicon Ultra-filters (UFC905008, Millipore Sigma) and purified by size-exclusion chromatography using a Sephacryl S-500 HR size-exclusion column. For quality control, the final saponin adjuvant was characterized with the Limus Amebocyte Lystae assay (QCL-1000, Lonza) for low endotoxin levels. The adjuvant concentration was determined using a cholesterol quantification kit (MAK043, Sigma-Aldrich).

Recombinant immunogen production

eOD-60mer was produced recombinantly as described previously (*18*). The eOD-60mer gene composed of eOD-GT8 monomer fused to lumazine synthase was synthesized by Integrated DNA Technologies, cloned into phLsec plasmid, and transfected into Expi293F (A14527, Thermo Fisher Scientific) cells. After 6 days, cell culture supernatant was harvested by centrifugation and filtered through a 0.2-μm filter. The eOD-60mer was affinity purified by incubating with *Galanthus nivalis* lectin–conjugated agarose beads (AL-1243-5, Vector Laboratories) overnight under gentle agitation at 4°C and eluted with lectin elution buffer containing 1 M methyl a-D-mannopyranoside (M6882, Millipore Sigma). The resulting solution was dialyzed in PBS and further purified by size-exclusion chromatography using Sephacryl S-500 HR resin.

N332-GT2 trimers were expressed in FreeStyle 293F cells (R79007, Invitrogen) and purified in two steps by affinity chromatography using a GE HisTrap column and size-exclusion chromatography using a GE S200 Increase column as described previously (*35*, *50*). N332-GT2 nanoparticles were expressed and purified as described previously (*35*).

Antigen labeling and characterization

Protein antigens (eOD-60mer, HIV Env trimer, trimer-NPs, and αCD35-eOD) at 1 mg/ml in PBS were diluted at a 1:1 volume ratio in 0.2 M sodium bicarbonate buffer (S8875, Sigma-Aldrich), pH 8.4, and kept on ice. Fresh 1 mg/ml stock solutions of Sulfo-Cyanine 3 (21320, Lumiprobe) and Sulfo-Cyanine 5 (23320, Lumiprobe) NHS esters were made in 0.2 M sodium bicarbonate, pH 8.4, and added to the antigen solutions. This reaction was allowed to proceed for 16 hours at 4°C, and then samples were desalted by passing through a Zeba Spin Desalting column (89882 and 87766, Thermo Fisher Scientific) and equilibrated in PBS twice. Labeled antigens were sterilized by filtering through 0.22-μm pore size Spin-X centrifuge tube filters (CLS8160, Millipore Sigma) and stored at 4°C until use. Antigen degree of labeling was determined by measuring the absorbance at 280, 568, and 646 nm wavelengths to measure total protein, Cy3 dye, and Cy5 dye,

respectively. To calculate the concentration of the different constituents, extinction coefficient values of 41200, 113215, 141390, 85550, 162000, and 271000 M^{-1} cm^{-1} were used for one subunit of eOD-60mer, N332-GT2 trimer, one subunit of N332-GT2 trimer-ferritin (24mer), αCD35-eOD, sulfo-cy3 NHS ester, and sulfo-cy5 NHS ester, respectively. The degree of labeling for either the nanoparticle or soluble antigens was calculated based on the ratio of antigen concentration to Cy3 or Cy5 concentration.

Complement-binding assay for labeled eOD-60mer

High-binding ELISA plates (07-200-37, Fisher Scientific) were coated with 1 μg/ml labeled eOD-$60mer_{40}$ and blocked with 2% bovine serum albumin (BSA) in PBS. Serum was collected from naïve mice using a collection tube with serum gel (41.1500.005, Sarstedt), diluted in PBS (3% v/v), and incubated in ELISA plates for 1 hour at 37°C. Complement binding was detected by biotinylated anti-C3 antibody (NB200-540B, Novus Biologicals), followed by streptavidin-horseradish peroxidase (HRP) (3310-9-1000, Mabtech AB). The plates were developed with 1-Step Ultra TMB-ELISA Substrate Solution (34028, Thermo Fisher Scientific), and the reaction was stopped by adding with 2 N sulfuric acid (BDH7500-1).

Antigenicity profiling of labeled immunogens

Antigenicity profiles of immunogens after dye conjugation were assessed through binding of structure-sensitive mAbs to plate-bound immunogens by ELISA. eOD-60mer with varying amounts of dyes conjugated were directly coated onto Corning High Binding plates (07-200-37, Fisher Scientific) at 2 μg/ml and blocked with 2% BSA (A8022, Sigma-Aldrich) dissolved in PBS. For trimers or trimer-NPs, plates were coated with 2 μg/ml *Galanthus nivalis* lectin (L8275, Sigma-Aldrich) before blocking and capture of 2 μg/ml antigen. Antigenicity profiles were evaluated by adding indicated human mAbs at titrated concentrations to plate-bound antigen, followed by detection with mouse antihuman secondary antibody conjugated to HRP (1721050, Bio-Rad Laboratories). HRP binding was determined based on reaction with 3,3′,5,5′-tetramethylbenzidine (TMB) substrate (34028, Thermo Fisher Scientific) and was stopped by the addition of 2 M sulfuric acid (BCH7500-1, VWR) at a 1:1 volume ratio. The optical density of the mixture was read out at 450 nm minus the absorbance at 540 nm according to the manufacturer's instructions.

In vitro eOD-$60mer_{40}$ digestion assays

Agarose bead–immobilized TPCK-treated trypsin (150 μl, 20230, Thermo Fisher Scientific) was prepared and equilibrated in 0.1 M ammonium bicarbonate buffer (S8875, Sigma-Aldrich) pH 8 according to the manufacturer's instructions. These beads were mixed with 100 μl of 20 μg/ml FRET dye-labeled Ags diluted in ammonium bicarbonate buffer and incubated at 37°C. The digested antigens were isolated at specified time points by centrifugation to remove the trypsin beads and characterized using three different assays to identify changes in molecular weight, *E*, and antigenicity. Molecular weight changes of the antigen were evaluated by reducing SDS-PAGE, and protein bands were visualized with high-sensitivity Flamingo Fluorescent Protein Gel Stain (1610491, Bio-Rad Laboratories). Intact antigen fractions were determined by digital imaging of gels followed by ImageJ analysis. *E* and antigenicity changes were obtained by directly coating the recovered antigens onto glass coverslips for imaging or plates for ELISA, respectively.

In addition to trypsin, eOD-$60mer_{40}$ was digested with 150 μg/ml recombinant MMP9 (909-MM-010, R&D Systems), 150 μg/ml MMP14 (918-MP-010, R&D Systems), or 100 μg/ml ADAM17 (2978-AD-010, R&D Systems) according to the manufacturers' instructions with minor modifications. For MMP9, 2.5 μM $ZnCl_2$ was added to the assay buffer and 1.5 mM p-aminophenylmercuric (APMA) was used to activate 150 μg/ml of MMP9. For MMP14, 150 μg/ml protease was activated by 3.5 μg/ml Furin (1503-FE, R&D Systems) for 2 hours at 37°C in activation buffer containing 50 mM Tris, 2.5 mM $CaCl_2$, and 0.25% w/v Brij-35 adjusted to pH 9. For 60mer digestion, 1 μM $ZnCl_2$ was added to the MMP14 activation buffer. For ADAM17, freshly reconstituted protease at a 100 μg/ml concentration and eOD-$60mer_{40}$ were both desalted in Zeba columns equilibrated with deionized water before mixing together to initiate cleavage. For all cases, 10 μg/ml eOD-$60mer_{40}$ was incubated with individual metalloproteinases or left in control conditions (APMA-containing assay buffer for MMP9, Furin-containing assay buffer for MMP14, or alone in deionized water for ADAM17) for 48 hours at 37°C. The digested 60mer was diluted to 1.5 μg/ml and analyzed on a TECAN Infinite 200Pro plate reader for solution-based FRET measurements (Ex/Em: 555/665 nm) and Cy5 (Ex/Em: 630/665 nm) fluorescence measurement or coated onto high-binding plates for VRC01 mAb (1 μg/ml)–binding ELISA.

Antigen stability within lymph and plasma

Lymph was collected from the mesenteric lymphatic duct as described previously (*51*) with a slightly modified terminal procedure. In brief, C57BL/6 mice received 200 μl of commercial olive oil by oral gavage 60 min before lymph collection to allow for easier visualization of the lymphatic ducts. At the time of collection, mice were anesthetized with isofluorane (4% for induction, 2.5% for maintenance) and their abdomens were shaved and prepared with ethanol and Betadine scrubs. A 3-cm incision was made in the skin and peritoneum along the midline. Intestine and colon were retracted to the side using sterile cotton swabs to expose the superior mesenteric artery and the adjacent mesenteric lymph ducts. The needle of a precision syringe (30 G, Hamilton) was inserted in one of the lymph ducts with the bevel pointing toward the distal portion of duct, and 5 to 10 μl of lymph was collected by gently retracting the plunger. For plasma collection, naïve C57BL/6 mice were bled retro-orbitally and processed in MiniCollect Tubes with EDTA (450480, MiniCollect).

To examine antigen stability, 10 μg/ml eOD-$60mer_{40}$ or $trimer_6$ was incubated in diluted plasma or lymph (10% v/v in PBS) or 150 μg/ml porcine trypsin (T4549, Millipore Sigma) at 37°C. The FRET (Ex/Em: 555/665 nm) and Cy5 (Ex/Em: 630/665 nm) signals were recorded at specified time points using a TECAN plate reader.

Antigen degradation after incubation with serum

To determine whether complement deposition onto antigen prevents enzymatic degradation, 20 μg/ml eOD-60mer was incubated for 15 min at 37°C with 10% fresh serum from naïve mice or left alone in PBS. This solution was mixed with trypsin-conjugated agarose beads and digested overnight at 37°C. Complement binding and structural integrity of the nanoparticles after digestion was assessed by ELISA using 1 μg/ml anti-C3 and human VRC01 mAb, respectively.

Mouse immunizations

All animal studies were performed under an institutional animal care and use committee–approved animal protocol following local, state, and National Institutes of Health guidelines for the care and use of animals. C57BL/6 mice were anesthetized and immunized with 10 μg of the indicated antigens in the presence or absence of 5 μg pf saponin adjuvant subcutaneously at the left and right sides of the tail base.

Tissue processing

Inguinal LNs extracted from euthanized mice were submerged into cryomolds containing optimal cutting temperature compound (23-730-571, Fisher Scientific) and dipped into 2-methylbutane (M32631, Millipore Sigma) prechilled in liquid nitrogen for 5 to 10 min. All frozen tissues were cryosectioned on a Leica CM1950 at 10-μm thickness, adhered to Superfrost Plus microscope slides (12-550-15, Fisher Scientific), and stored at −80°C until use.

Confocal microscopy

For all experiments, imaging was performed on a Leica SP8 confocal microscope equipped

with a white light laser and spectral emission filter to detect emission wavelengths between 470 and 670 nm with a minimum bandwidth of 10 nm. All images were recorded with a 25× water-immersion lens; for assessing antigen drainage in the LNs, laser settings were kept constant across different time points for each immunogen.

FRET imaging and analysis

To assess antigen integrity, an acceptor photobleaching method was used to compare Cy3 dye emission intensity before and after photobleaching the Cy5 dye. FRET imaging was performed on a Leica SP8 confocal microscope with a 25× objective, imaging selected square regions with dimensions ranging from 175 to 250 μm. Cy3 signal was recorded by exciting the dye at 555 nm and collecting its emission between 565 and 615 nm wavelengths, and Cy5 signal was recorded after exciting the dye at 640 nm and collecting emission between 660 and 720 nm wavelengths. To photobleach the Cy5 dye, regions of interest in LN sections were excited with 640-nm wavelength lasers at maximum power until the Cy5 intensity was <10% of its prebleached values; the duration of the bleaching process varied between 90 and 180 s. Cy3 and Cy5 fluorescence signals were then collected again after acceptor bleaching. E of antigens was calculated using a custom MATLAB code (*52*) based on established methods (*16*). E at each pixel (E_{ij}) is given by

$$E_{\mathrm{ij}} = 1 - \frac{(1-\alpha)Cy3_{\mathrm{Bef}}}{\gamma Cy3_{\mathrm{Aft}} - \alpha Cy3_{\mathrm{Bef}}}$$

where $Cy3_{\mathrm{Bef}}$ and $Cy3_{\mathrm{Aft}}$ are Cy3 emissions before and after Cy5 photobleaching. Correction factors α and γ account for incomplete bleaching of the Cy5 ($0 \le \alpha \le 1$) and unintended photobleaching of Cy3 ($\gamma \ge 1$), respectively. α values were experimentally determined from the ratio of Cy5 emission signals before and after photobleaching, and γ was determined to be ~1.05 on the basis of coverslip-coated antigens conjugated only with Cy3. Moreover, correction factors δ, the cross-talk from acceptor emission into the donor emission channel, and ε, cross-talk of photodegraded acceptor product into donor emission channel, which are listed in the full equation for E are found to be 0.004 ± 0.00094 and 0.0062 ± 0.00077, respectively, from monolayer experiments with antigens labeled only with Cy5. Thus, both δ and ε were approximated to be zero. To quantify E, antigen locations were first identified by using a binary mask generated from thresholding intensity values in the Cy5 image before photobleaching. To this end, nonimmunized LNs were imaged under same conditions to approximate the background fluorescence value of Cy5 needed for thresholding. This mask was applied to the remaining images to quantify $Cy3_{\mathrm{Bef}}$ from the Cy3 image before photobleaching, $Cy3_{\mathrm{Aft}}$ from the Cy3 image after photobleaching, and α from the Cy5 images before and after photobleaching. These parameters were used to generate a histogram of E values for antigens found within each imaging area. To determine the fraction of degraded antigen, this procedure was repeated for an intact control sample composed of FRET dye–conjugated antigen coated onto a glass coverslip. To quantify the amount of fully intact antigen at each time point/condition was determined by quantify the fraction of antigen$^+$ pixels in each imaged region with E_{ij} values larger than the minimal E value detected in the intact control sample.

For examining fractions of intact antigens colocalized with specified protein markers, the acceptor photobleaching procedure was performed on LN sections that were fixed, immunostained with antibodies, and mounted in PBS. In addition to the four images required for calculating E, fluorescence images of cell markers were recorded and overlaid onto regions of intact and degraded antigens.

Anti-CD45 antibody–conjugated eOD monomer for extracellular antigen stability analysis

Anti-CD45 mAb (103102, BioLegend) was conjugated to Cy5-labeled eOD-GT8 monomer with an added free N-terminal (*53*) cysteine using Sulfo-SMCC linker (A39268, Thermo Fisher Scientific). Sulfo-SMCC linker was first added to 1 mg/ml mAb solubilized in PBS at 20-fold molar excess. This reaction was performed for 45 min at room temperature and was terminated by removing unreacted linkers with PBS-equilibrated Zeba columns. This solution was temporarily maintained at 4°C.

The dye-labeled eOD monomer was next conjugated to Sulfo-SMCC tethered anti-CD45 mAb. To this end, 1 mg/ml of the Cy5-labeled eOD monomer, maintained in 1× Tris-buffered saline (BP24711, Fisher Scientific), was reduced by adding twofold molar excess of tris(2-carboxyethyl)phosphine (TCEP, 20490, Thermo Fisher Scientific). This reaction was stopped after 15 min at room temperature by using PBS-equilibrated Zeba desalting columns. The reduced eOD monomer was immediately mixed with Sulfo-SMCC linker tethered antibody at 1:1 mass ratio and was allowed to react overnight under mild agitation at 4°C. This mixture was repeatedly washed through Amicon ultracentrifugal units with 100-kDa MWCO (UFC510024, Millipore Sigma) to remove unreacted eOD monomer. The resulting antibody-eOD monomer construct was sterilized by passing through centrifugal filters with 0.2-μm pores (CLS8160-96EA, Millipore Sigma) and kept at 4°C.

For immunization of C57BL/6 mice, indicated doses of saponin adjuvant and antibody-eOD conjugate was administered subcutaneously at the tail base, and inguinal LNs were harvested and mechanically processed at specified time points. The isolated cells were first stained with Live/Dead Aqua (L34957, Thermo Fisher Scientific), anti-B220 BV421 antibody (103239, BioLegend), and anti-CD3ε antibody (100321, BioLegend) before splitting into two distinct pools.

To detect surface-bound or intact eOD monomer, each pool of cells was stained with either biotinylated anti-Cy3/Cy5 mAb (C3117, Millipore Sigma) or biotinylated VRC01 mAb, respectively, followed by exposure to streptavidin-BUV737 (612775, BD Biosciences) and analysis by flow cytometry.

Bulk antigen stability analysis using pull-down assay

Pairs of inguinal LNs from two C57BL/6 mice immunized with eOD-60mer$_{40}$ were harvested and pooled into prechilled tubes each containing 120 μl of 0.5% w/v BSA (A8022, Sigma-Aldrich), 2× Halt protease inhibitor cocktail (78430, Thermo Fisher Scientific), and 5 mM EDTA (15575020, Thermo Fisher Scientific) solubilized in PBS. The LNs were mechanically dissociated using a Biomasher II microtissue homogenizer (K7496250010, Fisher Scientific) and kept at 4°C for 30 min under gentle agitation and away from light. To obtain the extracellular protein fraction, the homogenized LN solution was centrifuged for 5 min at 500*g* before transferring 100 μl of the supernatant into a V-bottom 96-well plate (12-565-215, Thermo Fisher Scientific). In addition, 100 μl of intact antigen standards containing 300, 30, 3, or 0.3 ng of eOD-60mer$_{40}$ or eOD-60mer conjugated with only 20 Cy5 dyes (eOD-60mer$_{20\ Cy5}$) was added to the 96-well plate. To obtain the intracellular protein fraction from the LNs, 120 μl of cell lysis solution composed of T-PER (78510, Thermo Fisher Scientific), 2× Halt protease inhibitor cocktail, and 5 mM EDTA was added to the each of the centrifuged tubes with pelleted cells and tissue debris. These tubes were vortexed vigorously to dissociate the pellets and maintained under moderate agitation for 1 hour at room temperature and away from light. The tubes were then centrifuged at 1500*g* for 10 min, and 100 μl of the supernatant containing intracellular proteins were transferred to the V-bottom 96-well plate. For all wells containing LN samples and standards, 7.5 μl of 1 mg/ml biotinylated anti-Cy3/Cy5 antibody was added, and the plate was kept on a shaker overnight at 4°C.

The resulting antibody-antigen complexes were next captured onto microparticles. Specifically, 32.5 μl of streptavidin-conjugated magnetic particles (88816, Thermo Fisher Scientific) were first transferred to individual wells in a V-bottom 96-well plate and mixed with 170 μl of PBS. Using a 96-well side-skirted magnet (12027, Thermo Fisher Scientific), the beads

were pulled to the bottom of the wells and the supernatant was removed. The pelleted beads were resuspended in 105 μl of solution containing LN extracts or eOD-60mer standards bound to biotinylated anti-dye antibodies and kept on a shaker overnight at 4°C.

The beads bound with eOD-60mer$_{40}$ were next analyzed using the BD FACSymphony A3 flow cytometer in which the FRET signal (561 nm excitation, 670/30 nm emission) and acceptor signal (640 nm excitation, 670/30 nm emission) of the captured antigens were recorded. To quantify the mass of antigen within LNs, a standard curve was first generated based on the Cy5 mean fluorescence intensities from eOD-60mer$_{40}$ standards with known masses. Next, the raw mass values determined based on the standard curve was scaled by a correction factor to account for the inefficiency of eOD-60mer$_{40}$ retrieval from LN samples. Specifically, this factor (~1.8) was calculated by dividing the Cy5 signal from beads incubated with 3 ng of eOD-60mer$_{40}$ standard kept in BSA plus protease inhibitor solution with the Cy5 signal from the same standard mass mixed into mechanically dissociated naïve LN solution or lysate. The reported 60mer masses detected in LNs are the product of the correction factor and values obtained from the standard curve.

To evaluate the fraction of intact 60mer bound to the beads, normalized FRET (nFRET), defined as mean value of FRET divided by Cy5 intensity of all the particles, was first quantified for eOD-60mer$_{40}$ and eOD-60mer$_{20,Cy5}$ standards (made in PBS with BSA and inhibitors) that are used to represent 100% and 0% intact antigen, respectively. The nFRET value for eOD-60mer$_{40}$ is expected to be upper bound because all of the donor and acceptor dyes are present on the NP. By contrast, nFRET value for eOD-60mer$_{20,Cy5}$ serves as the lower bound because FRET does not occur without donor (Cy3) dyes. Moreover, nFRET was observed to vary with the amount of 60mer bound on the beads, so this value was quantified for beads incubated with serially diluted eOD-60mer$_{40}$ and eOD-60mer$_{20,Cy5}$ standards. The resulting data points were fitted to sigmoidal curves that provided the upper (100% intact) and lower (0% intact) limits of nFRET values for beads bound to any amount of 60mer. By assuming that the change in nFRET value between the upper and lower limits was linearly related to the percentage of intact antigen, the percentage of intact 60mer retrieved from LNs, the antigen amount of which was determined as described previously, was quantified.

Vibratome sectioning

Low-melting-point agarose (A4018, Sigma-Aldrich) solution was prepared by adding 0.4 g of agarose into 20 ml of PBS to achieve a concentration of 2% w/v. This mixture was microwaved at 30-s intervals until the agarose was fully dissolved, and the solution was equilibrated at 37°C for at least 45 min. Inguinal LNs were harvested from immunized or naïve mice and placed onto a 35-mm-diameter petri dish. The agarose solution was carefully added to this dish and incubated at 4°C for 10 min. A rectangular block encompassing the LNs was excised and mounted onto the sample holder of a Leica Vibratome VT1200s using Vetbond tissue adhesive (3M). The sample was held in PBS prechilled to 4°C and sectioned to obtain 250-μm-thick live tissue sections that were immediately transferred to a petri dish containing ice-cold RPMI 1640. All slices were collected and maintained in this solution until further use.

In vitro AZP assay

The sequences for the custom peptide (CPC Scientific) probes were as follows (*30*): quenched AZP, (5FAM)-GPVPLSLVMG-K(CPQ2)-(Peg2)-GC; D-isomer, UeeeeeeeeXGpvplslvmGrrrrrrrrrX-k(5-FAM)-amide; AZP, UeeeeeeeeXGPVPLSLVMGrrrrrrrrrX-k(5-FAM)-amide; and polyR, rrrrrrrrrX-k(Cy5)-amide, where uppercase letters indicate isomers, lowercase letters indicate D-isomers, 5-FAM is 5-carboxyfluorescein, Cy5 is cyanine 5, Peg2 is polyethylene glycol, and CPQ2 is the quencher.

To analyze the cleavage of peptide sequence in the AZP probe by different recombinant proteases, a quenched version of the probe that consisted of a peptide sequence flanked by quencher CPQ2 and fluorophore FAM was used. This reaction was performed by incubating 1 μM quenched AZP probe with 1 μg/ml recombinant MMP14 (918-MP-010, R&D Systems), ADAM17 (2978-AD-10, R&D Systems), ADAM10 (946-AD-020, R&D Systems), Cathepsin D (1029-AS-010, R&D Systems), Cathepsin S (50769-M08H, Sino Biologicals), or Cathepsin L (1515-CY-010, R&D Systems). The buffer pH for digestion was adjusted based on each manufacturer's protocol and were 8.5, 9.0, 9.0, 3.5, 4.5, and 6.0 for MMP14, ADAM17, ADAM10, Cathepsin D, Cathepsin S, and Cathepsin L, respectively. The FAM signal (Ex/Em: 485/535 nm) was recorded on a plate reader initially and after 60 min of incubation at 37°C.

For visualizing protease activity within tissues, LNs were harvested from naïve mice or mice immunized with eOD-60mer and saponin at the indicated time points. These tissues were embedded within 2% agarose gel containing 5 μM AZP and polyR control probes, vibratome sliced to obtain 250-μm-thick live sections, and maintained within RMPI 1640 supplemented with 2 μM $ZnCl_2$. Parallel tissue sections were also incubated in 5 μM uncleavable D-isomer probe and polyR as a control. After 2 hours of incubation at 37°C, the tissue-gel constructs were washed twice with PBS before fixation in 4% paraformaldehyde overnight at 4°C. Excess paraformaldehyde was removed by washing in PBS, and these sections were imaged directly by confocal microscopy. From the images, areas of positive AZP probe signal normalized to areas with positive polyR probe signal were calculated for inside versus outside of the follicles.

To immunophenotype proteolytically active cells, live LN vibratome sections were placed into RMPI medium containing the following combination of probes at 2 μM concentration: (i) AZP and polyR probes, (ii) D-isomer and polyR probe, and (iii) AZP probe only. These slices were maintained in the probe medium for 2 hours at 37°C before washing away the excess probe by performing two cycles of transferring the slices to and maintaining in fresh PBS for 15 min at 37°C. The LN sections were subsequently processed for flow cytometry analysis. To assess protease activity, different cellular phenotypes were first identified and AZP$^+$ cells were determined only among the polyR$^+$ cells within each cell type. Gating for polyR and AZP probe binding was determined based on cells extracted from slices incubated in probe media conditions (iii) and (ii), respectively (fig. S5A).

Ex vivo modulation of protease activity within LN slices

To examine the impact of proteases on FDC-localized antigens, mice were first immunized with saponin and eOD-60mer$_{40}$. After 3 days, LNs were harvested and vibratome sliced to 250 μm thickness and were left untreated in RPMI 1640 medium or exposed to either 200 μg/ml porcine trypsin (T4549, Millipore Sigma) or 10 μg/ml recombinant full-length murine MMP9 without additional chemical activation by p-aminophenylmercuric acetate (ab39309, Abcam). After 4 hours at 37°C, tissue slices were washed in PBS and fixed overnight in 10% neutral buffered formalin (HT5011-1CS, Sigma-Aldrich) at 4°C before confocal imaging.

Similarly, the effect of protease inhibitors on antigen degradation was assessed by immunizing mice with saponin and eOD-60mer$_{40}$, followed by harvesting of LNs 2 hours later for vibratome sectioning. These live slices were left untreated in RPMI 1640 medium or incubated with 100 μg/ml marimastat (M2699, Millipore Sigma) or Halt broad-spectrum protease inhibitor without EDTA (78430, Thermo Fisher Scientific) at a fivefold dilution of the stock solution. After 6 hours in 37°C, the tissue slices were washed in PBS and fixed overnight in 10% neutral buffered formalin at 4°C.

In both ex vivo experiments, the fixed sections were incubated in fresh PBS for 1 hour at 25°C, followed by flash freezing and cryosectioning. The sections were adhered to microscope slides, mounted in PBS, and imaged immediately.

Effect of in vivo metalloproteinase inhibition on antigen stability within LNs

Marimastat (Selleck Chemicals) was dissolved in dimethylsulfoxide (DMSO) to achieve a concentration of 52 mg/ml. This stock solution was diluted to 9 mg/ml in PBS, sonicated for 5 min, and injected intraperitoneally at a volume of 50 μl per mouse, resulting in an ~25 mg/kg dose. Control mice were administered matching concentrations of DMSO diluted in PBS. These injections were performed twice daily for 4 days. On the fifth day, marimastat was administered 2 hours before and 6 hours after immunization with 5 μg of saponin adjuvant and 10 μg of eOD-60mer$_{40}$. Draining LNs were extracted 24 hours after immunization for antigen pull-down assay and FRET imaging.

MMP9 transfer plasmid design

Gene fragments encoding the full-length human MMP9 with a G100L mutation (to achieve functional proteolysis without a requirement for propeptide cleavage) (*34*) and C-terminal HA affinity tag was purchased from Integrated DNA Technologies. This gene block was inserted into a MIGR1 retroviral transfer plasmid, followed by an IRES sequence and the gene encoding for mNeongreen reporter protein.

Retrovirus production

Phoenix Eco cells (CRL-3214, ATCC) were plated at a density of 75,000 cells/cm^2 in a 10-cm-diameter tissue culture dish and maintained in 10 ml of Dulbecco's modified Eagle's medium containing 10% fetal bovine serum (FBS), 2 mM L-glutamine, and 1% penicillin/streptomycin (complete medium) for 24 hours. To produce retroviruses, these cells were transfected with retroviral vectors using lipofectamine 3000 (L3000015, Thermo Fisher Scientific) according to the manufacturer's instructions. In brief, 1.5 ml of Opti-MEM (31985062, Thermo Fisher Scientific) containing 41 μl of lipofectamine 3000 reagent was mixed with 1.5 ml of Opti-MEM containing 35 μl of P3000 reagent, 4.5 μg of pCL-Eco packaging plasmid (12371, Addgene), and 13.5 μg of transfer plasmid containing constitutively active MMP9/mNeongreen or mCherry inserts. This mixture was maintained at room temperature for 20 min. Before transfection, the culture medium for Phoenix Eco cells was replaced with 6 ml of Opti-MEM and lipofectamine-plasmid mixture was added dropwise over the surface of the culture dish. The Phoenix Eco cells were transfected overnight before replacing the Opti-MEM with 10 ml of complete medium. The generated viruses were harvested after 48 hours by collecting the medium, centrifuging at 2000*g* for 5 min, and filtering through 0.45-μm polyvinylidene difluoride syringe filters (24-242, Genesee Scientific). The retrovirus was used immediately after or flash frozen in liquid nitrogen for extended storage.

Transduction of splenic B cells and adoptive transfer

Spleens were harvested from naïve C57BL/6 mice or CD45.1$^+$ C57BL/6 mice (strain 002014, The Jackson Laboratory), and B cells were isolated using a commercially available negative selection kit (19854, STEMCELL Technologies) according to the manufacturer's instructions. The isolated B cells were resuspended at a density of 1 million cells/ml in base medium, distributed at 2 ml per well in six-well plates, and activated for 24 hours with 10 μg/ml lipopolysaccharide (LPS, L2630-10MG, Millipore Sigma) and 25 ng/ml interleukin-4 (404-ML-025, R&D Systems). The base medium was composed of RPMI 1640, 10% v/v FBS, 55 μM β-mercaptoethanol (21-985-023, Fisher Scientific), and 1% v/v penicillin/streptomycin.

To transduce the B cells using retrovirus, 2 ml of 20 μg/ml RetroNectin (T100B, Takara Bio) was first coated onto each well of non-tissue culture–treated six-well plates (351146, Corning) on the same day that splenic B cells were isolated. The plates were kept overnight at 4°C before removing the RetroNectin solution and blocking with sterile 2% BSA solubilized in PBS for 30 min at room temperature. The plates were next washed once with PBS before adding 1.3 ml of supernatant with MMP9/mNeongreen- or mCherry-encoding retrovirus per well (see the "Retrovirus production" section) and centrifuging the plate at 2000*g* for 2 hours at 32°C to attach the virus onto RetroNectin-coated plates. Next, the supernatant was removed, wells were washed three times with PBS, and activated B cells were transferred to the wells to undergo transduction for 48 hours. The expression of MMP9 within the culture medium was quantified by using a commercially available ELISA kit (ab100610, Abcam). The activity of secreted MMP9 was validated by concentrating the culture medium by eightfold, transferring the supernatant onto anti-HA tag Ab (901533, BioLegend)–coated ELISA plates preblocked with 2% BSA in PBS, and exposing the captured MMP9 in the wells to 1 μM quenched AZP probe. The change in carboxyfluorescein (FAM) intensity was recorded using a plate reader as described previously.

For adoptive transfer, transduced B cells were collected from the wells and centrifuged, the supernatant was aspirated, and the pellet was resuspended in 15 ml of base medium. The washing process was repeated three more times to completely remove LPS, and after the final wash, the B cells were resuspended in PBS at a density of 2.5 million cells per 10 μl. This cell mixture was kept on ice. To perform intranodal injections into the inguinal LNs, mice were anesthetized and the area between the hind thigh and abdomen was shaved and depilated with a mild hair removal cream. Left and right superficial inguinal LNs were visualized with the help of a gooseneck fiber optic illuminator as defined opaque spots that are 5 to 7 mm superior and lateral to the corresponding inguinal (fourth) nipple. The LN was gently pushed up against the skin, which was pulled taut to allow for controlled insertion of the needle. The needle of a precision syringe (30 G, Hamilton) was inserted under the skin and pushed 1 to 2 mm into the LN along its long axis. For each LN, 10 μl of B cell suspension was slowly injected, and the entry of this solution into the tissue was confirmed by the visible swelling of the LN.

scRNA-seq analysis

The scRNA-seq data were collected and analyzed in a separate study. The data were used here to identify cell-type-specific expression of proteases. Briefly, scRNA-seq (10X Genomics) was performed according to the manufacturer's instructions on dissociated cells from LNs of mice that were either untreated or 6 hours after vaccination with 200 μg of ovalbumin peptide and 20 μg CpG 1826 adjuvant. Cell Ranger was used to process sequence data. Gene expression is log normalized for each cell such that the total unit for each cell sums to 10,000. Clusters of the major cell types were derived using the Louvain clustering algorithm and annotated based on known markers. Genes encoding extracellular or secreted proteases, including ADAMs, ADAMTSs, MMPs, and elastases, were included in the analysis. For FDCs, scRNA-seq analysis was performed in a similar manner based on published datasets (*28*, *29*).

Immunofluorescence staining

Frozen sections were retrieved from −80°C, quickly thawed, and incubated in 10% neutral buffered formalin solution for 8.5 min at 25°C. The sections were washed three times in PBS with a 10-min incubation time between each wash. Excess PBS was removed after the final wash before incubating the slides in blocking buffer composed of 2% BSA and 2% Triton X-100 in PBS. After 30 min, the blocking buffer was aspirated and the slides were stained in the following primary antibody solutions also made in blocking buffer for ~16 hours at 4°C: 1:75 anti-ADAM17 (NBP2-67179, Novus Biologicals), 1:75 anti-ADAM10 (NBP176973, Novus Biologicals), 1:75 anti-MMP14 (ab51074, Abcam), 1:75 anti-CD35 (740029, BD Biosciences), 1:75 anti-LYVE1 (53-0443-82, Thermo Fisher Scientific), and 1:75 anti-CD169 (142421, BioLegend). These slides were washed in PBS three times for 10 min, stained with 1:200 diluted secondary antibodies solution in blocking buffer (ab150063 Abcam, ab150061, Abcam) for 4 hours at room temperature, and washed again in PBS.

To mount the slides, 5 μl of PBS was added directly onto the stained tissues before gently placing a 20 × 20 mm coverslip on top of the PBS droplet to sandwich the section. The coverslip was sealed using CoverGrip coverslip sealant (23005, Biotium) and imaged immediately. Slides were stored at 4°C for a maximum of 1 week.

To visualize five distinct markers within the same LN, modifications were made to the above staining protocol. Cryosectioned LNs were first fixed and blocked with 10% v/v normal donkey serum and 0.1% w/v Triton X-100 constituted in PBS. The slides were then sequentially stained with the following antibodies and reagents at the listed dilutions for the indicated times in the following order: (i) anti-ADAM17 mAb (NBP2-67179, Novus Biologicals), 1:75, 1 hour; (ii) donkey anti-rabbit polyclonal Ab conjugated to BV421 (406410, BioLegend), 1:250, 1 hour; (iii) anti-ADAM10 mAb conjugated to AF647 (NBP2-12015AF647, Novus Biologicals), anti-MMP14 mAb conjugated to AF555 (ab302548, Abcam), anti-CD169 mAb conjugated to AF488 (142419, BioLegend), and biotinylated anti-LYVE1 mAb (13-0443-82, Thermo Fisher Scientific), all at 1:75; and (iv) streptavidin conjugated to BV510 (405233, BioLegend), 1:150, 0.5 hour.The slides were washed three times in PBS after each staining and mounted as described previously after streptavidin binding. The LNs were imaged on a Leica SP8 scanning confocal in which three distinct scan sequences were created to eliminate fluorophore cross-talk. For the first sequence, the LN section was excited with a 405-nm laser, and emission between 410 and 450 nm and 620 and 710 nm was recorded to acquire ADAM17 and LYVE1$^+$ LECs signal. For the second sequence, the tissue was excited with a 488- and a 650-nm laser, and emission between 498 and 530 nm and 660 and 770 nm were recorded to detect CD169$^+$ subcapsular macrophages (SSMs) and ADAM10, respectively. For the final sequence, the slide was excited with a 555-nm laser, and emission between 565 and 620 nm were obtained to visualize MMP14. The colocalization of the proteases with SSMs or LECs were quantified using the coloc2 plugin on ImageJ. Regions of interests were created on either SSM or LEC images and colocalization with each protease was represented using Mander's coefficient, which required images to have $P > 0.95$ after Costes analysis.

Follicle targeting anti-CD35 ScFv-eOD monomer construct

Follicle targeting of antigens was achieved by using an anti-CD35 mAb (clone: 8C12) that specifically recognized complement receptor 1 (CR1), which is highly expressed by FDCs. To express this antibody efficiently, the framework sequences within the variable region of both light and heavy chains of the original clone, 8C12 (*54*), were mutated toward rat germline amino acid sequences using the SabPred toolbox (*55*). This construct was expressed as a single-chain variable fragment (ScFv) fused to eOD monomer (αCD35-eOD). eOD monomer fused to anti-GFP ScFv was used as a control construct (Iso-eOD). Gene fragments encoding these fusion proteins and flanked with overhangs designed for insertion into NotI restriction enzyme linearized GWIZ vector were purchased from Integrated DNA Technologies. Plasmid assembly and protein expression was performed as described previously.

αCD35-eOD

HHHHHHGGDTITLPCRPAPPPHCSSNITGLILTRQGGYSNDNTVIFRPSGGDWRDIARCQIAGTVVSTQLFLNGSLAEEEVVIRSEDWRDNAKSICVQLNTSVEINCTGAGHCNISRAKWNNTLKQIASKLREQYGNKTIIFKPSSGGDPEFVNHSFNCGGEFFYCDSTQLFNSTWFNSTGGGGSGGGGSGGGGS**DIVMTQTPSSLAVSAGEKVTMSCKSSQSLLYSKNKKNYLAWYQQKPGQSPKLLISWASSRESGVPDRFTGSGSGTDFTLTISSVQAEDLAVYYCEQYYNIPYTFGGGTKLELKRGGGGSGGGGSGGGGSQVKLQESGGGLVQPGRSLKLSCAASGFTFSNYDMAWVRQAPTKGLEWVASINYDGSSTYYRDSVKGRFTISRDNAKSTLYLQMDSLRSEDTATYYCTTLYNWYVMDAWGQGTTVTVSS**

Iso-eOD

HHHHHHGGDTITLPCRPAPPPHCSSNITGLILTRQGGYSNDNTVIFRPSGGDWRDIARCQIAGTVVSTQLFLNGSLAEEEVVIRSEDWRDNAKSICVQLNTSVEINCTGAGHCNISRAKWNNTLKQIASKLREQYGNKTIIFKPSSGGDPEFVNHSFNCGGEFFYCDSTQLFNSTWFNSTGGGGSGGGGSGGGGS**DIQLTQSPAIMSPSLGERVTMTCTASSSVGSSYLHWFQQKPGSSPKLWIYSTSNLASGVPARFSGSGSGTSYSLTISRMEAEDAATYYCHQYHRTPYTFGGGTKLEIKRAGGGGSGGGGSGGGGSEVQLQQSGPELVKPGSSMKISCKASGYSFTGYTMNWVKQSHGQNLEWIGLINPYNGGTNYNQKFKGKATLTVDKSSSTAYMELLGLTSEDSAVYYCTRGNSDFSAWFAYWGQGTSVTVSS** Where underlining indicates the eOD monomer and bold indicates the ScFv construct.

Human tonsil tissue acquisition and processing

Adult human tonsil samples sourced from standard surgical procedures performed for clinical care was provided by the Massachusetts Eye and Ear Infirmary at the Massachusetts General Hospital. Usage for scientific research was approved by the Massachusetts General Hospital institutional review board under protocol P2020-003521. Tonsil samples were deidentified and processed within 2 to 3 h after removal. Acquired tonsil tissues were flash frozen, cryosectioned, and immunostained in the same manner as murine LNs.

Flow cytometry

For analysis of AZP binding and protease expression, inguinal LNs from immunized or naïve mice or live LN sections were incubated in 1 ml of digestion solution composed of RPMI 1640 medium containing 0.8 U/ml dispase (07913, STEMCELL Technologies), 0.1 mg/ml Collagenase D (11088858001, Sigma-Aldrich), and 0.1 mg/ml of DNAse (10104159001, Sigma-Aldrich). After 20 min of incubation at 37°C, the LNs were repeatedly passed through a 1-ml pipette tip to mechanically rupture the tissue. Care was taken to collect ~1 ml of the solution containing only the liberated cells before transferring to ice-cold PBS containing 5 mM EDTA and 2% FBS (stop buffer). The remaining tissue debris were further processed by a second incubation with 1 ml of fresh digest buffer for 20 min at 37°C and transfer of the supernatant to stop buffer. This solution was passed through a 40-μm pore size filter, centrifuged, and transferred into a V-bottom 96-well plate for flow staining. For GC analysis, cells were isolated by mechanically crushing the LNs against a filter cap with 70-μm pore size and generously pipetting cold PBS containing 1% BSA and 5 mM EDTA (FACS buffer) onto the porous mesh. The cells were centrifuged, resuspended in flow cytometry buffer, passed through a 40-μm pore size filter, and transferred to a V-bottom 96-well plate.

The isolated cells were stained with Live/Dead fixable aqua stain (L34957, Thermo Fisher Scientific) for 10 min at 25°C before washing twice in flow cytometry buffer. Cells were then incubated with Fc block for 10 min at 4°C before staining with antibodies listed in the key resource table (see the supplementary materials) for 20 additional min at 4°C. For on-target trimer and off-target probe specific GC B cell analysis, cells stained with antibodies were distributed evenly and exposed to biotinylated trimer or peptide probes preincubated with streptavidin (30 min at molar ratio of 1:4 at 25°C) conjugated to phycoerythrin (405203, BioLegend), BV421 (405226, BioLegend), or allophycocyanin (40520L7, BioLegend). For AZP analysis, the stained cells were immediately analyzed (*28*). For protease expression analysis, cells stained with surface markers were fixed and permeabilized using a transcription factor staining buffer kit (00-5523-00, Invitrogen), followed by incubation with antiprotease antibody for 30 min at 25°C. After washing twice with wash buffer, the cells were incubated with donkey antirabbit antibody conjugated to Alexa Fluor 488 dye to detect the proteases. For compensation, UltraComp

eBeads Plus compensation beads (01-3333-42, Thermo Fisher Scientific) were used to bind to dye-labeled antibodies. Stained cells were analyzed using a BD FACSymphony A3 cell analyzer.

Probe synthesis for characterizing trimer breakdown product–specific GC B cells

Peptides corresponding to and buried within BG505 MD39 were synthesized by GenScript. Sequences of the peptides were as follows: HR1 peptide-biotin, GIKQLQARVLAVEHY-LGAGK-biotin; C1 peptide-biotin, DQSLKPA-VKLTPLGAGK-biotin; and V3 peptide-biotin, TRPNNNTVKSIRIGPGQAFYYTGGAGK-biotin.

ELISA analysis of serum antibody responses

Blood collected from immunized mice through retro-orbital bleeding was processed in collection tubes with separator gel (41.1500.005, Sarstedt) to obtain serum samples. To analyze on-target antibody response, high-binding ELISA plates (07-200-37, Fisher Scientific) were coated with 1 μg/ml trimer and blocked with 2% BSA in PBS. For off-target responses, plates were first coated with 2 μg/ml streptavidin (434302, Thermo Fisher Scientific), then blocked with 2% BSA in PBS, and finally incubated with 2 μg/ml of HR1, C1, or V3 peptides diluted in blocking buffer. To detect antigen-specific IgG responses, dilutions of serum were added to the wells and incubated for 1.5 hours at 25°C. Plates were washed three times in PBS containing 0.2% Tween-20, and then anti-mouse IgG secondary antibody conjugated to HRP (172-1011, Bio-Rad Laboratories), diluted 1:5000 in blocking buffer as per manufacturer instructions, was added to the wells. After 1 hour of incubation, plates were again washed, developed with TMB, and stopped with sulfuric acid. End-point titers are reported as inverse dilutions in which absorbance at 450 nm minus the reference absorbance at 540 nm equals 0.1.

Quantification and statistical analysis

Information pertaining to statistical analysis such as n (number of mice or tissue slices or image regions) and P or q values are indicated in the figure legends. All plots within the figures show means with SD. No methods were used to determine whether data met assumptions of statistical approach. GraphPad Prism 9 was used to calculate statistical significance from tests specified within each figure legend. All results are from at least two independent experiments. Unless specified otherwise, significance from statistical tests within the figures are denoted by asterisks in the figures: $*P \leq 0.05$, $**P \leq 0.01$, $***P \leq 0.001$, and $****P \leq 0.0001$; nonsignificant results are represented by n.s.

REFERENCES AND NOTES

1. A. J. Pollard, E. M. Bijker, Publisher correction: A guide to vaccinology: from basic principles to new developments. *Nat. Rev. Immunol.* **21**, 129 (2021). doi: 10.1038/s41577-020-00497-5; pmid: 33402728
2. M. J. Shlomchik, W. Luo, F. Weisel, Linking signaling and selection in the germinal center. *Immunol. Rev.* **288**, 49–63 (2019). doi: 10.1111/imr.12744; pmid: 30874353
3. G. D. Victora, M. C. Nussenzweig, Germinal centers. *Annu. Rev. Immunol.* **30**, 429–457 (2012). doi: 10.1146/annurev-immunol-020711-075032; pmid: 22224772
4. R. K. Abbott *et al*., Precursor frequency and affinity determine B cell competitive fitness in germinal centers, tested with germline-targeting HIV vaccine immunogens. *Immunity* **48**, 133–146.e6 (2018). doi: 10.1016/j.immuni.2017.11.023; pmid: 29287996
5. O. Bannard, J. G. Cyster, Germinal centers: Programmed for affinity maturation and antibody diversification. *Curr. Opin. Immunol.* **45**, 21–30 (2017). doi: 10.1016/j.coi.2016.12.004; pmid: 28088708
6. P. Dosenovic *et al*., Anti-HIV-1 B cell responses are dependent on B cell precursor frequency and antigen-binding affinity. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4743–4748 (2018). doi: 10.1073/pnas.1803457115; pmid: 29666227
7. M. Kuraoka *et al*., Complex antigens drive permissive clonal selection in germinal centers. *Immunity* **44**, 542–552 (2016). doi: 10.1016/j.immuni.2016.02.010; pmid: 26948373
8. H. A. McNamara *et al*., Antibody feedback limits the expansion of B cell responses to malaria vaccination but drives diversification of the humoral response. *Cell Host Microbe* **28**, 572–585.e7 (2020). doi: 10.1016/j.chom.2020.07.001; pmid: 32697938
9. K. M. Cirelli, S. Crotty, Germinal center enhancement by extended antigen availability. *Curr. Opin. Immunol.* **47**, 64–69 (2017). doi: 10.1016/j.coi.2017.06.008; pmid: 28738289
10. H. H. Tam *et al*., Sustained antigen availability during germinal center initiation enhances antibody responses to vaccination. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E6639–E6648 (2016). doi: 10.1073/pnas.1606050113; pmid: 27702895
11. B. S. Graham, M. S. A. Gilman, J. S. McLellan, Structure-based vaccine antigen design. *Annu. Rev. Med.* **70**, 91–104 (2019). doi: 10.1146/annurev-med-121217-094234; pmid: 30691364
12. D. M. Catron, K. A. Pape, B. T. Fife, N. van Rooijen, M. K. Jenkins, A protease-dependent mechanism for initiating T-dependent B cell responses to large particulate antigens. *J. Immunol.* **184**, 3609–3617 (2010). doi: 10.4049/jimmunol.1000077; pmid: 20208013
13. J. G. Tew, T. E. Mandel, A. W. Burgess, Retention of intact HSA for prolonged periods in the popliteal lymph nodes of specifically immunized mice. *Cell. Immunol.* **45**, 207–212 (1979). doi: 10.1016/0008-8749(79)90378-2; pmid: 88271
14. B. A. Smith *et al*., Persistence of infectious HIV on follicular dendritic cells. *J. Immunol.* **166**, 690–696 (2001). doi: 10.4049/jimmunol.166.1.690; pmid: 11123354
15. B. A. Heesters *et al*., Endocytosis and recycling of immune complexes by follicular dendritic cells enhances B cell antigen binding and activation. *Immunity* **38**, 1164–1175 (2013). doi: 10.1016/j.immuni.2013.02.023; pmid: 23770227
16. J. Roszik, J. Szöllosi, G. Vereb, AccPbFRET: An ImageJ plugin for semi-automatic, fully corrected analysis of acceptor photobleaching FRET images. *BMC Bioinformatics* **9**, 346 (2008). doi: 10.1186/1471-2105-9-346; pmid: 18713453
17. P. Dosenovic *et al*., Immunization for HIV-1 broadly neutralizing antibodies in human Ig knockin mice. *Cell* **161**, 1505–1515 (2015). doi: 10.1016/j.cell.2015.06.003; pmid: 26091035
18. J. Jardine *et al*., Rational HIV immunogen design to target specific germline B cell receptors. *Science* **340**, 711–716 (2013). doi: 10.1126/science.1234150; pmid: 23539181
19. J. G. Jardine *et al*., HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen. *Science* **351**, 1458–1463 (2016). doi: 10.1126/science.aad9195; pmid: 27013733
20. J. G. Jardine *et al*., HIV-1 VACCINES. Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen. *Science* **349**, 156–161 (2015). doi: 10.1126/science.aac5894; pmid: 26089355
21. D. Sok *et al*., Priming HIV-1 broadly neutralizing antibody precursors in human Ig loci transgenic mice. *Science* **353**, 1557–1560 (2016). doi: 10.1126/science.aah3945; pmid: 27608668
22. T. Tokatlian *et al*., Innate immune recognition of glycans targets HIV nanoparticle immunogens to germinal centers. *Science* **363**, 649–654 (2019). doi: 10.1126/science.aat9120; pmid: 30573546
23. K. V. Houser *et al*., Safety and immunogenicity of a ferritin nanoparticle H2 influenza vaccine in healthy adults: A phase 1 trial. *Nat. Med.* **28**, 383–391 (2022). doi: 10.1038/s41591-021-01660-8; pmid: 35115706
24. B. J. Read *et al*., Mannose-binding lectin and complement mediate follicular localization and enhanced immunogenicity of diverse protein nanoparticle immunogens. *Cell Rep.* **38**, 110217 (2022). doi: 10.1016/j.celrep.2021.110217; pmid: 35021101
25. D. Y. Mason, M. Jones, C. C. Goodnow, Development and follicular localization of tolerant B lymphocytes in lysozyme/anti-lysozyme IgM/IgD transgenic mice. *Int. Immunol.* **4**, 163–175 (1992). doi: 10.1093/intimm/4.2.163; pmid: 1622894
26. L. Tang *et al*., Enhancing T cell therapy through TCR-signaling-responsive nanoparticle drug delivery. *Nat. Biotechnol.* **36**, 707–716 (2018). doi: 10.1038/nbt.4181; pmid: 29985479
27. M. M. Schmidt, G. M. Thurber, K. D. Wittrup, Kinetics of anti-carcinoembryonic antigen antibody internalization: Effects of affinity, bivalency, and stability. *Cancer Immunol. Immunother.* **57**, 1879–1890 (2008). doi: 10.1007/s00262-008-0518-1; pmid: 18408925
28. V. N. Kapoor *et al*., Gremlin 1$^+$ fibroblastic niche maintains dendritic cell homeostasis in lymphoid tissues. *Nat. Immunol.* **22**, 571–585 (2021). doi: 10.1038/s41590-021-00920-6; pmid: 33903764
29. L. B. Rodda *et al*., Single-cell RNA sequencing of lymph node stromal cells reveals niche-associated heterogeneity. *Immunity* **48**, 1014–1028.e6 (2018). doi: 10.1016/j.immuni.2018.04.006; pmid: 29752062
30. A. P. Soleimany *et al*., Activatable zymography probes enable in situ localization of protease dysregulation in cancer. *Cancer Res.* **81**, 213–224 (2021). doi: 10.1158/0008-5472.CAN-20-2410; pmid: 33106334
31. T. Jakoš, A. Pišlar, A. Jewett, J. Kos, Cysteine cathepsins in tumor-associated immune cells. *Front. Immunol.* **10**, 2037 (2019). doi: 10.3389/fimmu.2019.02037; pmid: 31555270
32. T. Yadati, T. Houben, A. Bitorina, R. Shiri-Sverdlov, The ins and outs of cathepsins: Physiological function and role in disease management. *Cells* **9**, 1679 (2020). doi: 10.3390/cells9071679; pmid: 32668602
33. A. P. Amini *et al*., Multiscale profiling of protease activity in cancer. *Nat. Commun.* **13**, 5745 (2022). doi: 10.1038/s41467-022-32988-5; pmid: 36192379
34. K. E. Fisher *et al*., Engineering autoactivating forms of matrix metalloproteinase-9 and expression of the active enzyme in cultured cells and transgenic mouse brain. *Biochemistry* **41**, 8289–8297 (2002). doi: 10.1021/bi012076t; pmid: 12081477
35. J. M. Steichen *et al*., A generalized HIV vaccine design strategy for priming of broadly neutralizing antibody responses. *Science* **366**, eaax4380 (2019). doi: 10.1126/science.aax4380; pmid: 31672916
36. J. T. Martin *et al*., Combined PET and whole-tissue imaging of lymphatic-targeting vaccines in non-human primates. *Biomaterials* **275**, 120868 (2021). doi: 10.1016/j.biomaterials.2021.120868; pmid: 34091299
37. K. M. Cirelli *et al*., Slow delivery immunization enhances HIV neutralizing antibody and germinal center responses via modulation of immunodominance. *Cell* **180**, 206 (2020). doi: 10.1016/j.cell.2019.12.027; pmid: 31923396
38. T. Ng'uni, C. Chasara, Z. M. Ndhlovu, Major scientific hurdles in HIV vaccine development: Historical perspective and future directions. *Front. Immunol.* **11**, 590780 (2020). doi: 10.3389/fimmu.2020.590780; pmid: 33193428
39. W. Kastenmüller, P. Torabi-Parizi, N. Subramanian, T. Lämmermann, R. N. Germain, A spatially-organized multicellular innate immune response in lymph nodes limits systemic pathogen spread. *Cell* **150**, 1235–1248 (2012). doi: 10.1016/j.cell.2012.07.021; pmid: 22980983
40. A. J. Dabo, N. Cummins, E. Eden, P. Geraghty, Matrix metalloproteinase 9 exerts antiviral activity against respiratory syncytial virus. *PLOS ONE* **10**, e0135970 (2015). doi: 10.1371/journal.pone.0135970; pmid: 26284919
41. J. P. Fernandes *et al*., Breast tumor-associated metalloproteases restrict reovirus oncolysis by cleaving the sigma1 cell attachment protein and can be overcome by mutation of σ1. *J. Virol.* **93**, e01380-19 (2019). doi: 10.1128/JVI.01380-19; pmid: 31462562
42. K. C. Hansen, A. D'Alessandro, C. C. Clement, L. Santambrogio, Lymph formation, composition and circulation: A proteomics perspective. *Int. Immunol.* **27**, 219–227 (2015). doi: 10.1093/intimm/dxv012; pmid: 25788586
43. C. C. Clement *et al*., Protein expression profiles of human lymph and plasma mapped by 2D-DIGE and 1D SDS-PAGE coupled with nanoLC-ESI-MS/MS bottom-up proteomics. *J. Proteomics* **78**, 172–187 (2013). doi: 10.1016/j.jprot.2012.11.013; pmid: 23202415

44. B. A. Tamburini, M. A. Burchill, R. M. Kedl, Antigen capture and archiving by lymphatic endothelial cells following vaccination or viral infection. *Nat. Commun.* **5**, 3989 (2014). doi: 10.1038/ncomms4989; pmid: 24905362
45. M. F. Bachmann, B. Odermatt, H. Hengartner, R. M. Zinkernagel, Induction of long-lived germinal centers associated with persisting antigen after viral infection. *J. Exp. Med.* **183**, 2259–2269 (1996). doi: 10.1084/jem.183.5.2259; pmid: 8642335
46. J. G. Cyster, B cell follicles and antigen encounters of the third kind. *Nat. Immunol.* **11**, 989–996 (2010). doi: 10.1038/ni.1946; pmid: 20959804
47. X. Y. Wang, B. Wang, Y. M. Wen, From therapeutic antibodies to immune complex vaccines. *NPJ Vaccines* **4**, 2 (2019). doi: 10.1038/s41541-018-0095-z; pmid: 30675393
48. P. W. Dempsey, M. E. Allison, S. Akkaraju, C. C. Goodnow, D. T. Fearon, C3d of complement as a molecular adjuvant: Bridging innate and acquired immunity. *Science* **271**, 348–350 (1996). doi: 10.1126/science.271.5247.348; pmid: 8553069
49. N. Pardi *et al*., Expression kinetics of nucleoside-modified mRNA delivered in lipid nanoparticles to mice by various routes. *J. Control. Release* **217**, 345–351 (2015). doi: 10.1016/j.jconrel.2015.08.007; pmid: 26264835
50. J. M. Steichen *et al*., HIV vaccine design to target germline precursors of glycan-dependent broadly neutralizing antibodies. *Immunity* **45**, 483–496 (2016). doi: 10.1016/j.immuni.2016.08.016; pmid: 27617678
51. B. Banan *et al*., Intestinal lymph collection via cannulation of the mesenteric lymphatic duct in mice. *J. Surg. Res.* **260**, 399–408 (2021). doi: 10.1016/j.jss.2020.11.004; pmid: 33261855
52. A. Aung, FRET efficiency calculation for acceptor photobleaching approach, Zenodo (2022); doi: 10.5281/zenodo.7327790
53. T. J. Moyer *et al*., Author correction: A Engineered immunogen binding to alum adjuvant enhances humoral immunity. *Nat. Med.* **26**, 804 (2020). doi: 10.1038/s41591-020-0861-0; pmid: 32405059
54. W. C. Lima *et al*., The ABCD database: A repository for chemically defined antibodies. *Nucleic Acids Res.* **48** (D1), D261–D264 (2020). doi: 10.1093/nar/gkz714; pmid: 31410491
55. J. Dunbar *et al*., SAbPred: A structure-based antibody prediction server. *Nucleic Acids Res.* **44**, W474–W478 (2016). doi: 10.1093/nar/gkw361; pmid: 27131379

ACKNOWLEDGMENTS

We thank the Koch Institute Swanson Biotechnology Center's Flow Cytometry and Microscopy core facilities for technical support; D. Lingwood for the generously providing the Hemagglutinin-Ferritin nanoparticle antigen; and W. du Fresne von Hohenesche of Geneva Antibody Facility for providing insights on improving the expression efficiency of the follicle-targeting antibody construct. **Funding:** D.J.I. received financial support from the Scripps Consortium for HIV/AIDS Vaccine Development (grant UM1 AI144462); the National Institutes of Health (NIH grant P01AI048240); the Ragon Institute of MGH, MIT, and Harvard; the Marble Center for Cancer Nanomedicine; and the Howard Hughes Medical Institute. S.N.B. was supported by the Koch Institute (support grant P30-CA14051), the National Institute of Environmental Health Sciences (grant P30-ES002109), the Marble Center for Cancer Nanomedicine, and the Howard Hughes Medical Institute. W.R.S. was supported by the Scripps Consortium for HIV/AIDS Vaccine Development (grant UM1 AI144462), the Scripps Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (grant UM1 AI100663), the Bill and Melinda Gates Foundation Collaboration for AIDS Vaccine Discovery (grants NAC INV-007522 and INV-008813), and the IAVI Neutralizing Antibody Center at TSRI. A.P.S. was supported by a NIH molecular biophysics training grant and a National Science Foundation graduate research fellowship. J.D.K. was supported by a Ludwig Center at MIT's Koch Institute fellowship. D.S.K. received financial support from the Ragon Institute of MGH, MIT, and Harvard. A.A. was supported by the NIH (training grant 5T32AI007386).
Author contributions: A.A. and D.J.I. conceived the study. A.A. and D.J.I. designed all of the experiments and methods, excluding RNA-seq, with additional contributions from L.M. for adoptive transfer experiments and lymph extraction along with A.P.A., J.D.K., and S.N.B. for AZP assays. A.C. and N.H. designed, performed, and analyzed scRNA-seq for LN cells. A.A. performed all other experiments with substantial contributions from J.R.G., M.B., L.M., H.L., M.S., H.S., P.A., T.R., S.X., W.A., and J.A. A.A. performed formal analysis of data with additional contributions from H.L. for image analysis and from Y.Z., D.M.M., and J.C.L. for FDC scRNA-seq analysis. A.P.A., J.D.K., and S.N.B. provided AZPs, C.A.C. and W.R.S. provided immunogens and antigen probes, and P.H., D.S.K., and L.M.F. provided human tonsils. A.A. and D.J.I. wrote the manuscript and prepared the figures. D.J.I. supervised the study. **Competing interests:** S.N.B. reports compensation for cofounding, consulting for, and/or board membership in Glympse Bio, Satellite Bio, CEND Therapeutics, Catalio Capital, Intergalactic Therapeutics, Port Therapeutics, Vertex Pharmaceuticals, and Moderna, and receives sponsored research funding from Johnson & Johnson, Revitope, and Owlstone. W.R.S. is an inventor on patents filed by Scripps and IAVI on the eOD-GT8 60mer, eOD-GT8 monomer, and BG505 MD39 trimer immunogens. N.H. holds equity in BioNtech and consults for Related Sciences/Danger Bio. D.J.I. and A.A. are inventors on a patent application related to CD35-targeted immunogens. The remaining authors declare no competing interests. **Data and materials availability:** All data are available in the main text, the supplementary materials, or have been deposited at Zenodo (*52*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn8934
Figs. S1 to S12
Data Files S1 and S2
MDAR Reproducibility Checklist

Submitted 29 December 2021; resubmitted 28 September 2022
Accepted 14 December 2022
10.1126/science.abn8934

Your Legacy to Science

AN ESTATE GIFT TO THE
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE





Since 1848, our founding year, the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you become a member of our 1848 Society, joining Thomas Edison, Alexander Graham Bell and the many distinguished individuals whose vision led to the creation of AAAS and our world-renowned journal, *Science*, so many years ago.

Unlike many of its peers, *Science* is not for-profit. Your estate gift would provide long-term financial stability and durable annual income that will support operations and competitive innovation for years to come. **This support is vital.**

"As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all."

—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

Legal Name: American Association for the Advancement of Science
Federal Tax ID Number: 53-0196568
Address: 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at **www.aaas.org/1848Society**.

1848 SOCIETY AAAS

cut here

Yes, I would like more information about joining the AAAS 1848 Society.

PLEASE CONTACT ME AT:

Name: _____

Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____

RETURN THIS FORM TO:
AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

RESEARCH ARTICLES

RESPIRATORY COMPLEXES

Structural basis of mammalian respiratory complex I inhibition by medicinal biguanides

Hannah R. Bridges[1]*, James N. Blaza[1,2], Zhan Yin[1], Injae Chung[1], Michael N. Pollak[3], Judy Hirst[1]*

The molecular mode of action of biguanides, including the drug metformin, which is widely used in the treatment of diabetes, is incompletely characterized. Here, we define the inhibitory drug-target interaction(s) of a model biguanide with mammalian respiratory complex I by combining cryo–electron microscopy and enzyme kinetics. We interpret these data to explain the selectivity of biguanide binding to different enzyme states. The primary inhibitory site is in an amphipathic region of the quinone-binding channel, and an additional binding site is in a pocket on the intermembrane-space side of the enzyme. An independent local chaotropic interaction, not previously described for any drug, displaces a portion of a key helix in the membrane domain. Our data provide a structural basis for biguanide action and enable the rational design of medicinal biguanides.

The biguanide metformin is central to the treatment of millions of patients with type 2 diabetes worldwide (*1*) and has been studied intensely in recent years for treatment of other conditions, including ischemia-reperfusion injury (*2*, *3*), fibrosis (*4*), viral infections (*5*), and cancer (*6*). Optimization of biguanides for distinctive indications has been hindered by incomplete understanding of their molecular pharmacology, and although preclinical evidence for the antineoplastic action of metformin was sufficient to justify dozens of clinical trials (*7*), the results have been disappointing (*8*, *9*). Biguanides have been reported to target many cellular proteins, including mitochondrial glycerophosphate dehydrogenase (mGPD) (*10*), presenilin enhancer 2 (*11*), F_1F_O adenosine 5′-triphosphate (ATP) synthase (*12*), cytochrome *c* oxidase (*13*), and the chloride intracellular channel 1 (*14*). Several studies have described biguanide inhibition of mitochondrial respiratory complex I (proton-translocating NADH:ubiquinone oxidoreductase, where NADH is the reduced form of NAD^+) (*12*, *15*–*18*), supporting a mode of biguanide action in which decreased production of ATP from oxidative phosphorylation triggers the activation of adenosine 5′-monophosphate kinase and inhibition of adenylate cyclase, leading to beneficial downstream effects on gluconeogenic enzymes (in diabetes) and mTOR (in cancer and antiviral treatments) (*1*, *6*, *18*, *19*).

Complex I is a 1-MDa multiprotein assembly that is central to mitochondrial and cellular metabolism. It oxidizes the NADH produced by oxidation of carbohydrates and lipids to maintain the redox state of the mitochondrial nicotinamide adenine dinucleotide (NAD^+) pool, reduces ubiquinone-10 to drive the respiratory chain and oxygen consumption, and pumps protons out of the mitochondrial matrix. This proton pumping contributes to the proton-motive force (PMF) that drives ATP synthesis through oxidative phosphorylation (*20*). Cryo–electron microscopy (cryo-EM) studies of complex I have revolutionized our understanding of its structure, mechanism, and regulation, informing on redox catalysis in the hydrophilic domain, proton translocation across the membrane, and possible mechanisms of coupling between ubiquinone-10 reduction and proton translocation (*21*, *22*). Furthermore, cryo-EM has discriminated different resting states of the enzyme (*22*–*25*) on the basis of domain-level reorientations linked to altered conformational states of the quinone-binding channel (Q-channel): the "active" state with a structurally ordered, turnover-ready Q-channel and the pronounced "deactive" state with a locally disordered Q-channel that requires restructuring and reactivation for catalysis. Biguanides bind with an unusual preference to the deactive state of the enzyme (*16*), but their binding site(s) and modes of interaction are unknown: They are expected to bind in a site downstream of the Fe-S clusters (*12*) but do not inhibit in a simple competitive manner. Their interaction site is expected to be amphipathic on the basis of the biguanide positive charge and strong correlations between inhibitory potency, cytotoxicity, and hydrophobicity (*12*, *26*).

We use cryo-EM to reveal the molecular interactions of biguanides with mammalian respiratory complex I, defining how they inhibit catalysis. We identify distinctive binding modes and rationalize biguanide protein-state selectivity to enable future implementation of structure-based drug design in the development of biguanide-based therapies for diverse applications.

[1]MRC Mitochondrial Biology Unit, University of Cambridge, The Keith Peters Building, Cambridge Biomedical Campus, Cambridge CB2 0XY, UK. [2]Structural Biology Laboratory and York Biomedical Research Institute, Department of Chemistry, The University of York, York YO10 5DD, UK. [3]Lady Davis Institute of the Jewish General Hospital and Department of Oncology, McGill University, Montreal, QC H3T 1E2, Canada.
*Corresponding author. Email: hrb@mrc-mbu.cam.ac.uk (H.R.B.); jh@mrc-mbu.cam.ac.uk (J.H.)

IM1761092 as a model biguanide for structural studies

The antidiabetic biguanides metformin and phenformin are relatively weak inhibitors of complex I, with simple molecular shapes. We sought to avoid technical risks in cryo-EM (excessive adventitious binding and reduced image contrast) from using these compounds in high millimolar concentrations by identifying a stronger-binding derivative, which would also exhibit a more distinctive cryo-EM density. For structural investigation of the complex I binding site(s) of biguanides, we therefore assessed IM1761092 (hereafter IM1092) (*27*), a more hydrophobic (log*P* 2.37) derivative of the metformin-related antidiabetic biguanide phenformin (log*P* 0.34) that contains a 3-chloro-4-iodo-phenyl ring (Fig. 1A and fig. S1). IM1092 inhibits cellular oxygen consumption (fig. S1D) and exhibits a stronger inhibition of complex I catalysis in bovine heart mitochondrial membranes than phenformin and metformin [half-maximal inhibitory concentration (IC_{50}) in the membrane is more than 10 and 2000 times lower, respectively, depending on the conditions; Fig. 1, B and C]. Its behavior is thus consistent with the reported correlation between biguanide inhibitory potency (IC_{50}) and hydrophobicity (log*P*) (*28*).

Mammalian mitochondrial membranes "as-prepared" contain a mixture of active and deactive complex I. In the deactive state, an important structural feature of the Q-channel, the loop between TMH1 and -2 in subunit ND3 (ND3 TMH1-2 loop) that carries Cys^{39} is disordered (*24*, *25*), but in the active state, it is ordered and Cys^{39} is buried (*23*, *25*, *29*, *30*). The two states can be discriminated biochemically by their sensitivity to *N*-ethyl maleimide (NEM), which derivatizes ND3-Cys^{39} in the deactive state, preventing catalysis, but leaves the active state unaffected. We found biguanide inhibition depends on the amount of the deactive state present in membranes (Fig. 1B). As-prepared membranes incubated with both 200 μM IM1092 (10 × IC_{50}) and NEM at 4°C exhibited essentially the same deactive content as biguanide-free controls (fig. S2A), indicating that biguanides do not shift the deactive:active population equilibrium in this condition. Furthermore, biguanide inhibition is stronger at higher pH (Fig. 1C). As biguanides [pK_a ~11 (where K_a is the acid dissociation constant) (*31*)] remain singly protonated at all pH levels tested, the pH dependence likely arises from changes in the protein, such as in local charges on residue side chains or phospholipid headgroups, or conformational changes. Having established that IM1092 inhibits cellular

oxygen consumption and shares the same selectivity for inhibiting catalysis by binding to the deactive state of complex I that is observed for antidiabetic biguanides, as well as the same pH-dependent mode of action, we proceeded with this tighter-binding synthetic biguanide as a representative model compound for structural studies for the technical reasons outlined previously.

To investigate the selectivity of biguanides for different states, IM1092 was added to the mixed population of states in the purified resting enzyme, without purposeful deactivation, for cryo-EM. No substrates were added because inhibition is weaker when biguanides are added during catalysis, rather than before (*12*). IM1092 has an IC_{50} value of 86 μM for catalysis of detergent-solubilized, purified bovine complex I (fig. S2B), and cryo-EM grid conditions with 350 μM IM1092 were chosen to maximize binding within the limits of pH- and biguanide-induced protein aggregation at high concentrations (fig. S2). When purified complex I (fig. S3) was incubated in the same conditions as for cryo-EM grid preparation and then diluted into inhibitor-free assay buffer, 98% of the control catalytic rate was recovered (fig. S2F), demonstrating full reversibility of inhibition under this condition.

Overview of cryo-EM particle populations

A total of 17,203 micrographs collected from a single cryo-EM grid yielded 598,287 good particles, which refined to an estimated global resolution of 2.1 Å (fig. S4), but with such a high degree of heterogeneity that it was not possible to model parts of the map. Subsequent global classification yielded three major classes resembling the active and deactive states mentioned

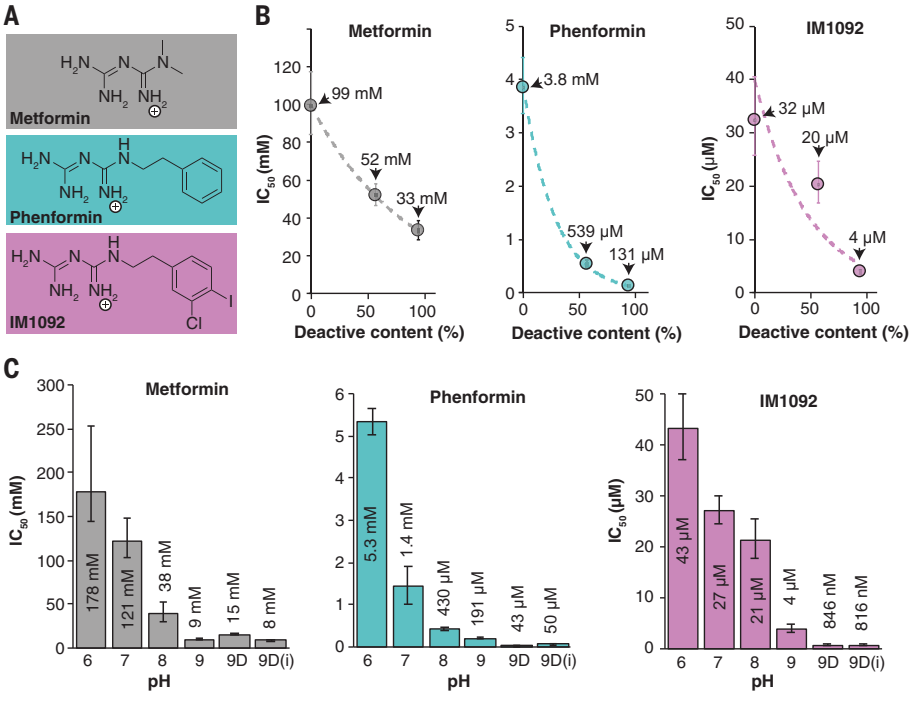


Fig. 1. Characterization of biguanide effects on catalysis and regions of structural interest. (**A**) Chemical structures of metformin, phenformin, and IM1092 in monoprotonated form. (**B**) Correlation between membrane deactive complex I content and IC_{50} for metformin (gray), phenformin (teal), and IM1092 (orchid). Error bars represent SEM for deactive content and 95% confidence intervals for IC_{50}. Data are fit to an exponential regression for visualization. (**C**) Effect of pH on IC_{50} in bovine heart membranes for metformin (gray), phenformin (teal), and IM1092 (orchid). 9D, pH 9 deactivated membranes; 9D(i), pH 9 deactivated membranes measured in the presence of antimycin A to inhibit complex III and the alternative oxidase to oxidize quinol by a different route, confirming inhibition is on complex I. Error bars represent 95% confidence intervals.

Table 1. Key characteristics of classes presented in this work. Unclear, unclear binding position, but confident identity; unidentified, the presence of density of unknown nonprotein origin.

Model	Classification scheme	PDB ID	Features		
			Q-channel ($C\text{-}C_{mask}$)	ND5 and NDUFB4 interface	ND2, NDUFB5, and NDUFA11 ($C\text{-}C_{mask}$)
Active inhibitor-free	Inhibitor-free (fig. S5)	7QSD	No density	Intact	NDUFC2 N terminus
Active-1092-i	Q-channel (fig. S7)	7R41	Unidentified	Mixed	IM1092-bound (0.49)
Active-1092-ii	Lateral helix (fig. S8)	7R42	Unidentified	Intact	IM1092-bound (0.52)
Active-1092-iii	Lateral helix (fig. S8)	7R43	Unidentified	Displaced	Unidentified
Active-1092-iv	Lateral helix (fig. S8)	7R44	Unidentified	Disordered	IM1092-bound (0.63)
Deactive-1092-i	Q-channel (fig. S7)	7R45	IM1092-bound (0.66)	Mixed	IM1092-bound (0.60)
Deactive-1092-ii	Q-channel (fig. S7)	7R46	IM1092-bound (0.72)	Mixed	Unidentified
Deactive-1092-iii	Q-channel (fig. S7)	7R47	IM1092-bound (0.68)	Mixed	IM1092-bound (0.68)
Deactive-1092-iv	Lateral helix (fig. S8)	7R48	IM1092-bound (0.66)	Intact	IM1092-bound (0.65)
Deactive-1092-v	Lateral helix (fig. S8)	7R4C	IM1092-bound (0.71)	Displaced	IM1092-bound (0.64)
Deactive-1092-vi	Lateral helix (fig. S8)	7R4D	IM1092 unclear	Disordered	IM1092-bound (0.67)
Slack-1092-i	Q-channel (fig. S7)	7R4F	IM1092 unclear	Mixed	IM1092-bound (0.65)
Slack-1092-ii	Q-channel (fig. S7)	7R4G	IM1092-bound (0.64)	Mixed	IM1092-bound (0.64)

above and a state called slack (supplementary text), which is of unknown functional relevance but has been described previously in cryo-EM studies of bovine complex I (*22*, *24*, *25*). Inhibitor-free reference maps of bovine complex I from separate inhibitor-free preparations in detergent [figs. S5 and S6 and EMD-3731 (*24*)] were used for comparison to identify and evaluate features found in the maps with IM1092 present. Three regions of interest are (i) densities occupying the Q-channel; (ii) unusually poor density in a portion of the C-terminal lateral helix in subunit ND5 and the adjacent subunit NDUFB4; and (iii) altered density at the expected position of the NDUFC2 subunit N terminus. Different classification strategies were tested to disentangle the different states and IM1092 interactions, leading to implementation of a "local-first" classification regime (supplementary text) used in two separate schemes to describe 12 distinct classes (Table 1 and figs. S7 to S9).

Biguanide-binding site 1: Ubiquinone-binding channel

Density was observed in the Q-channel, close to its entrance from the membrane, in all three major classes. Using local-first classification (fig. S7), we separated the particles into one active class (Active-1092-i; fig. S10), three deactive classes (Deactive-1092-i, -ii, and -iii; figs. S11 to S13), and two slack classes (Slack-1092-i and -ii; Table 1 and figs. S14 and S15). The major classes were assigned by global comparisons to reference maps and key local features (table S1 and supplementary text). Densities for IM1092 in the Q-channel are clear in the Deactive-1092-i, -ii, and -iii and Slack-1092-ii maps (Fig. 2, C to E and H, and figs. S16 and S17), and the cross-correlation fits (C-C_{mask}) (*32*) for the fit of the IM1092 molecule into its density were high (0.66 to 0.72) (Table 1). In Slack-1092-i, although the density is consistent for the chloro-iodo-phenyl moiety, the biguanide moiety of the putative IM1092 molecule was insufficiently resolved to confidently model its orientation (Fig. 2G). The density in the Active-1092-i map occupies a position similar to that of the densities observed in the other classes (Fig. 2F), but a smaller additional density feature is also observed further into the Q-channel so that, together, they resemble density observed in the Q-channel of the active-apo (inhibitor-free) class of bovine complex I in nanodiscs (EMD-14133) (*22*). With an IM1092 molecule refined in different orientations into the density in the Active-1092-i map near the Q-channel entrance, C-C_{mask} values are low (<0.5), and the shape of the density was visibly a poor fit; the second density was too small and featureless to ascertain its origin, and the identity of these two densities remain unconfirmed. Overall, ~45 to ~60% of the total population and ~56 to ~75% of the (deactive and slack) population presents clear evidence of IM1092 occupying the Q-channel.

In the deactive and slack states, IM1092 binds in an amphipathic site straddling two zones of the Q-channel: the hydrophobic region next to the exit (Fig. 2A) and the charged central region of the channel (*33*). The chloro-iodo-phenyl group of IM1092 points toward the channel exit (Fig. 2A), forming weak halogen bonds from the Cl to the ND1-Pro^{48} carbonyl and from the I to the S of ND1-Met^{225}, as well as van der Waals interactions with ND1-Phe^{224}, ND1-Phe^{220}, ND1-Leu^{55}, and NDUFS7-Trp^{46}. The halogen bonds are specific to IM1092, relative to phenformin, likely contributing to its higher binding affinity; and metformin, which lacks the phenyl ring, is not stabilized by the above interactions or by π-stacking to ND1-Phe^{224}, consistent with its much lower binding affinity. The biguanide moiety faces into the channel, toward the charged region, and adopts a range of different orientations (Fig. 2, C to E and H). In the Deactive-1092-i and Slack-1092-ii states, it forms a cation-π interaction with NDUFS7-Trp^{46} and, in the Deactive-1092-i state, a weak ionic interaction with the ND1-Glu^{24} carboxyl also. In the Deactive-1092-ii and -iii states, its orientation brings it closer

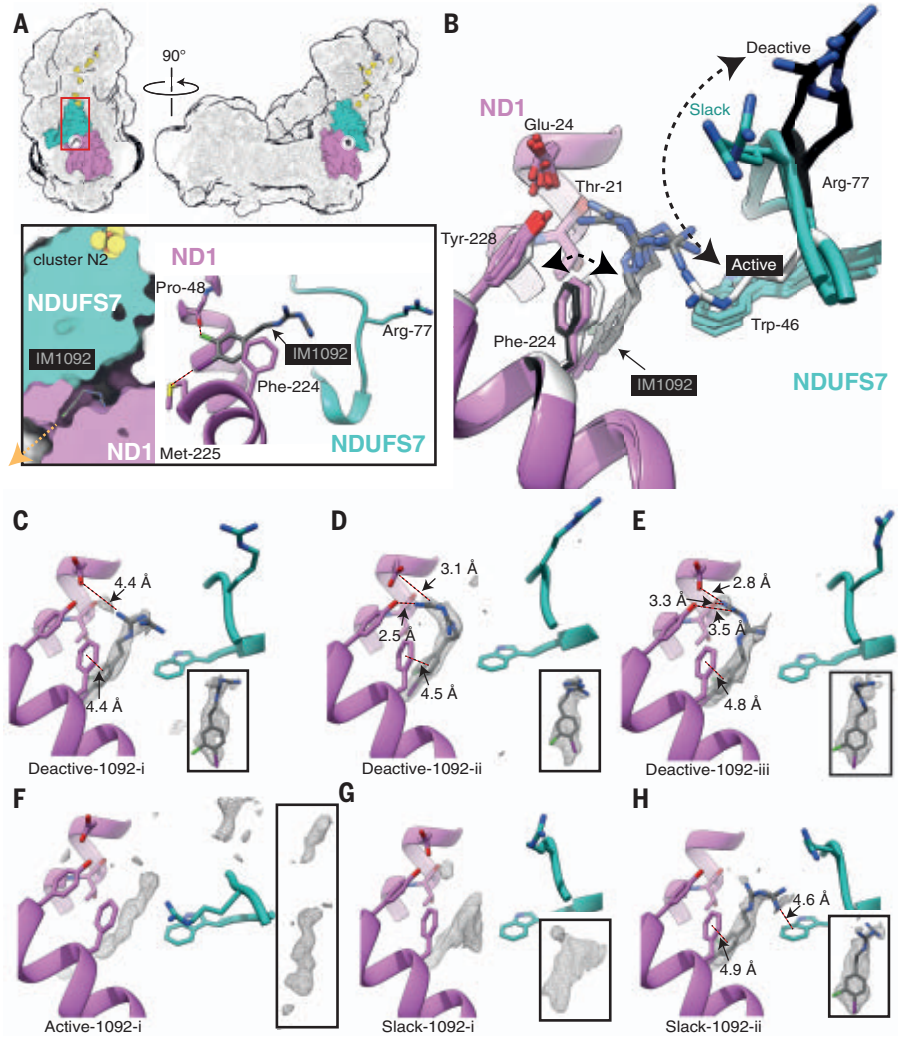


Fig. 2. Binding of IM1092 in the Q-channel. (**A**) Overview showing the location of the biguanide-binding site and inset showing a closer view and bonding interactions of the chloro-iodo-phenyl group for Deactive-1092-i. Orange arrow shows the route for exit from the Q-channel into the lipid bilayer. (**B**) Overlay of models for the Active-1092-i; Deactive-1092-i, -ii, and -iii; and Slack-1092-i and -ii states, aligned to subunit ND1, showing the location and variability of biguanide-binding orientations and relative position of NDUFS7-Arg^{77}. NDUFS7-Arg^{77} and ND1-Phe^{224} are white in the active models, mint or orchid in the slack models, and black in the deactive models. (**C** to **H**) Cryo-EM difference map densities (composite versus models) for biguanides bound to Deactive-1092-I (C), Deactive-1092-ii (D), Deactive-1092-iii (E), Active-1092-i (F), Slack-1092-i (G), and Slack-1092-ii (H). The insets show the difference map density for the biguanide for each model shown. Biguanides are not modeled in (F) and (G), owing to uncertainties in the ligand identity or orientation. Side-chain and ligand density for (C) to (H) are shown in fig. S16.

to ND1-Glu[24], and it forms a hydrogen bond with the ND1-Tyr[228] hydroxyl. In each state with modeled biguanide, the experimental IM1092 map density at higher thresholds is consistent with a mixture of inhibitor binding poses.

Notably, the conformation of the biguanide-binding region of the Q-channel differs between the three major states (*21*–*23*), regardless of whether an inhibitor is bound (Fig. 2B). Notably, NDUFS7-Arg[77] swings in an arc from its position in the active state, with its side chain pointing toward the Q-channel exit, to point approximately toward cluster N2 in the deactive state, adopting an intermediate position in the slack state (Fig. 2B). The NDUFS7-Arg[77] guanidinium is ~7 Å (slack) and 10 to 14 Å (deactive) away from the bound biguanide (N-N distance), but only ~4 Å away when IM1092 is refined into the density in the Active-1092-i map in the same orientation, indicating that repulsion between their positive charges may disfavor biguanide binding in the active state.

Available Protein Data Bank (PDB) models for mammalian, fungal, bacterial, and plant complexes I were compared to assess the conservation of key residues interacting with the biguanide moiety; NDUFS7-Trp[46], ND1-Glu[24], ND1-Tyr[228], and NDUFS7-Arg[77] are conserved in all species surveyed. The only high-quality reported human mutation in these residues in the ClinVar database (*34*) is the ND1-E24K (Glu[24]→Lys) mutation associated with Leber hereditary optic neuropathy–mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes (LHON-MELAS) overlap syndrome (*35*). Human mutations of the residues involved in the biguanide-binding site are therefore rare, consistent with their important role in enzyme function. Because they are mitochondrially encoded and/or essential for function, these residues cannot easily be artificially mutated in a mammalian system. Methods for mutating mitochondrial DNA are advancing, but site specificity and the efficiency of reaching homoplasmic edits remain challenging (*36*).

Biguanide interaction site 2: ND5 lateral helix and NDUFB4 interface

Transmembrane subunit ND5 contains an unusual long helix that runs laterally alongside ND4 and ND2, which has been proposed to either stabilize the proton-pumping modules or act as a transmissive element in proton pumping (*37*). Substantial evidence of disorder at the ND5 lateral helix–subunit NDUFB4 interface was observed in preliminary maps, so the dataset was subject to a separate local-first classification that focused on this region (figs. S8 and S18). The strategy yielded three major classes: one with a typical well-ordered ND5 lateral helix and NDUFB4 and two with distortion or disordering of ND5 residues 547 to 564 and nearby NDUFB4 residues 77 to 92. The classes were further separated into Active-1092-ii, -iii, and -iv (figs. S19 to S21); Deactive-1092-iv, -v, and -vi (Table 1 and figs. S22 to S24); and three slack classes, which all displayed poor density for the downstream ND5 lateral helix and were not further investigated. Overall, ~65% of the protein population was perturbed in this region, with similar proportions observed for active, deactive, and slack (~52, ~66, and ~58%, respectively). Deactive classes with a disrupted helix exhibit a small "opening" of the angles between the membrane and hydrophilic domains and distal and proximal membrane domains compared with their better-ordered equivalents (fig. S25 and supplementary text).

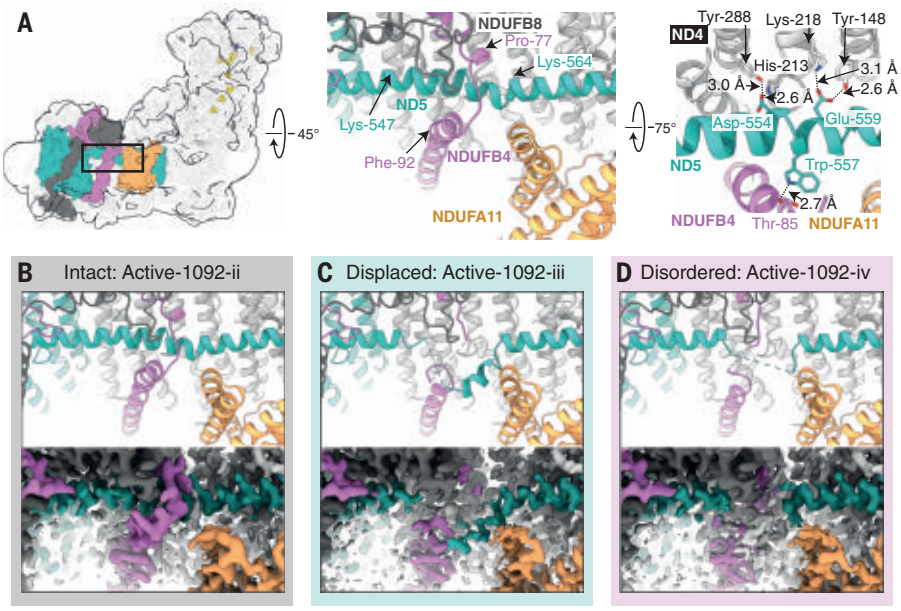


Fig. 3. Biguanide-induced distortion and disordering of the ND5 lateral helix and NDUFB4 loop. (**A**) Location of the structure disturbance and two views of the model of Deactive-1092-iv with hydrogen-bonding interactions indicated with black dotted lines and distances indicated. Orchid, NDUFB4; teal, ND5; orange, NDUFA11; dark gray, NDUFB8. Side-chain density for this region of Deactive-1092-iv is shown in fig. S18. (**B** to **D**) Models and composite cryo-EM maps for Active-1092-ii (B), Active-1092-iii (C), and Active-1092-iv (D) showing progressive disordering within the series. Details of π-bulge stabilizing interactions and equivalent disordering for the deactive states (Deactive-1092-iv, -v, and -vi) are shown in fig. S18.

This interesting region of the usually well-ordered ND5 lateral helix does not form a perfect α helix even in inhibitor-free active or deactive mammalian enzyme structures (*21*–*25*). The Active-1092-ii and Deactive-1092-iv models match the well-ordered inhibitor-free active model in this region. Two π-bulges (fig. S18) are stabilized by interactions between the lateral helix and nearby waters and by ionic and hydrogen bonding to ND4 (Fig. 3A, right).

In Active-1092-iii and Deactive-1092-v (referred to as displaced), a short portion of the ND5 lateral helix is altered: Lys[547] to Ser[550] are disordered, and an interruption of the helical structure at Leu[562] to Pro[563] allows a short stretch of helix (residues 550 to 559) to move outwards, away from the complex, and laterally, along the membrane plane (Fig. 3 and fig. S18). A loop in NDUFB4 at residues 78 to 83, which usually wraps around the lateral helix, becomes disordered from Pro[77] to Leu[91]. In Active-1092-iv and Deactive-1092-vi (referred to as disordered), the whole ND5 region from Lys[547] to Lys[564] appears disordered, along with the NDUFB4 loop described previously (Fig. 3D and fig. S18). Although not observed in the detergent-solubilized protein, the bovine enzyme in nanodiscs (*22*) contains two phospholipids near to the distorted region. No nearby density features can be interpreted as a tightly bound biguanide. Considering the positive charge on IM1092, it may interact with ND5-Asp[554] and/or Glu[559], thereby disrupting hydrogen bonding from the lateral helix to subunit ND4, or interact with nearby stabilizing phospholipids. Generally, π-bulges are energetically unfavorable elements (*38*) that require stabilization by hydrogen bonding to polar side chains or water molecules (*39*). The strained nature of this region may make it particularly prone to destabilization by guanidium-like biguanides. No mutations of Asp[554] or Glu[559] are observed in the ClinVar database (*34*), and acidic residues in these positions are conserved in current mammalian and plant structures, which suggests their important role in stabilizing the membrane domain.

Biguanide-binding site 3: ND2, NDUFB5, and NDUFA11

Density that matches IM1092 was observed in a pocket formed by subunits ND2, NDUFB5, and NDUFA11 on the intermembrane-space

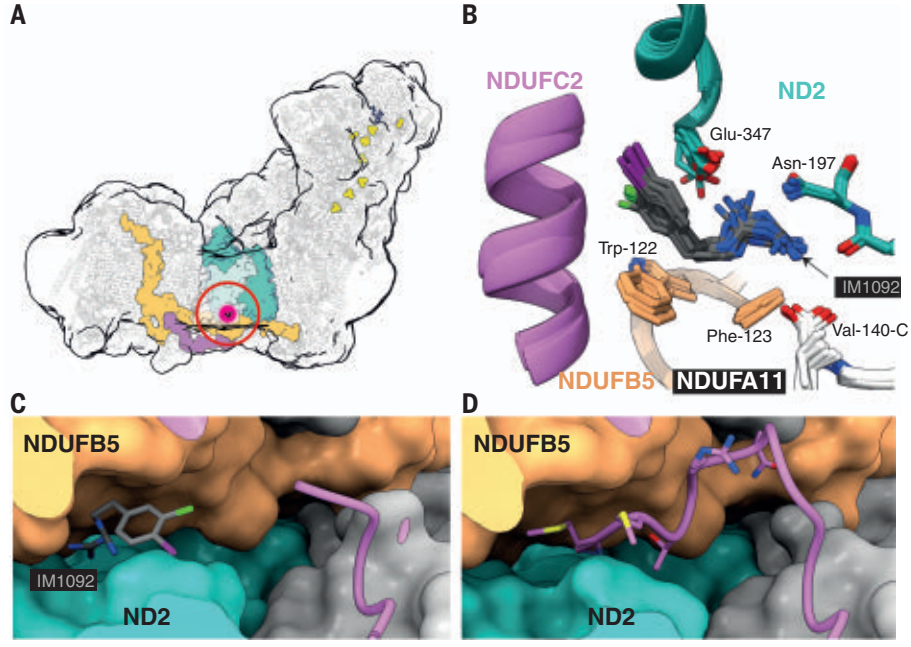


Fig. 4. Location of the biguanide binding at the ND2-NDUFB5-NDUFA11 interface. (**A**) Overall location of the binding site. Purple, NDUFC2; teal, ND2; orange, NDUFB5. (**B**) Overlay of all models aligned to ND2. White, NDUFA11; Val-140-C, C terminus of NDUFA11. (**C**) Surface representation (Active-1092-ii model) showing subunit NDUFC2 in cartoon, with atoms shown for IM1092. (**D**) Same view from the active inhibitor-free model showing the atoms from residues 1 to 7 of NDUFC2. Local interactions and density for the individual maps and models are shown in figs. S26 and S27.

side of the enzyme (Fig. 4 and figs. S26 and S27). The pocket is occupied by the N-terminal seven residues of NDUFC2 in the inhibitor-free enzyme, both in n-dodecyl-B-D-maltoside (DDM) and in all states of the bovine enzyme in nanodiscs (*22*). The NDUFC2 N terminus is displaced by the biguanide, with inhibitor densities observed in most classes in this study (figs. S26 and S27). The biguanide moiety is stabilized by an ionic interaction with ND2-Glu^{347} and by hydrogen bonds with the ND2-Tyr^{196} backbone carbonyl, ND2-Thr^{199} hydroxyl, ND2-Asn^{197} side chain, and NDUFA11-Val^{140} C-terminal carboxyl, as well as nearby water molecules resolved in some maps (figs. S26 and S27). The IM1092 bound in this site does not interfere with any known catalytically relevant structural elements in the complex and so is likely to represent a noninhibitory interaction.

Independence of interaction sites

The two local classification schemes used (figs. S7 and S8) each yielded six models, and a summary of their key features is shown in Table 1. All six classes from Q-channel classification (fig. S7) had poor density for the ND5-NDUFB4 interface, consistent with a mixture of the three lateral helix states being represented there. Overall, ~65% of the imaged protein population is disrupted in this region (fig. S8), and disorder is observed in all three major classes, as well as in states with (e.g., Deactive-1092-i) or without clear binding of IM1092 at the Q-channel (Active-1092-i). Furthermore, in the classes from the lateral helix classification (fig. S8), density for IM1092 was observed in the Q-channel of deactive classes with both ordered and displaced lateral helix. Taken together, biguanide binding in the Q-channel and the state of the lateral helix are not correlated. IM1092 is observed in the ND2, NDUFB5, and NDUFA11 pocket regardless of the occupancy of the Q-channel or the status of the lateral helix (Table 1). Therefore, the three interaction sites are independent of one other.

Discussion

Biguanide access to complex I binding sites

Substantial inhibition of complex I in vivo requires the biguanide positive charge, an inherent feature of the two cojoined guanidinium moieties at physiological pH (*31*), to drive biguanide accumulation in the mitochondrial matrix (*40*) by up to 1000-fold relative to the cytosol in response to the mitochondrial PMF. This mitochondrial concentrating effect makes the intramitochondrial biguanide concentration sufficient for the inhibition of targets with only relatively weak biguanide affinity, such as complex I, making them as potentially relevant as targets outside of the mitochondrial matrix that exhibit greater intrinsic affinities, such as rat mGPD (*10*). Recent work has questioned mGPD itself as a therapeutic target because metformin was found to be noninhibitory against the human enzyme (*41*). All three interaction or binding sites described in complex I contain acidic residues close to the membrane-aqueous interface, suggesting that IM1092 accesses them from the membrane, likely with the chloro-iodo-phenyl group acting as an anchor into the hydrophobic membrane core and the hydrophilic biguanide moiety interacting with the negatively charged phosphate headgroups, as proposed previously (*42*, *43*). The most likely route of access for hydrophobic biguanides to the Q-channel is therefore via the matrix-facing phospholipid leaflet. Although the Q-channel is reproducibly well ordered in the active state, deactive (or slack) states have mobile regions of the Q-channel loops that face the mitochondrial matrix (*21*–*24*, *44*), so it is also possible that the Q-channel may become exposed to the matrix in these states, providing an alternative route for hydrophilic biguanides with poor membrane solubility, such as metformin, to enter the Q-channel.

Major inhibitory site and selectivity for the deactive state

Biguanides, including IM1092 and the antidiabetic compounds metformin and phenformin, all share a preference to bind to the deactive state of complex I. This behavior is not observed for any other class of complex I inhibitor, and we expect all three biguanides to share the same inhibitory binding site and mode of action. In comparison with canonical hydrophobic complex I inhibitors such as rotenone, biguanides are relatively hydrophilic, water-soluble molecules that are unlikely to bind in highly hydrophobic or membrane-intrinsic sites. We suggested previously that metformin might interact at the junction of the hydrophilic and hydrophobic domains where a set of mobile elements (in subunits NDUFS7, NDUFS2, ND3, and ND1) change their conformation between the active and deactive states by binding to a resting state and preventing a return to catalysis (*12*). Our structures now demonstrate that the biguanide-binding site is inside the Q-channel, in a region that likely becomes exposed to the matrix in the deactive state, and adjacent to the mobile element in NDUFS7 that carries Arg^{77} and that switches its conformation between the active and deactive states.

In the Q-channel, IM1092 binds in an amphipathic region with the biguanide moiety stabilized by hydrogen bonding, cation-π interactions, and ionic bonding, and the hydrophobic chloro-iodo-phenyl group is stabilized by weak halogen bonding and van der Waals interactions. These additional stabilizing interactions between the hydrophobic portion of the Q-binding site and hydrophobic biguanides explain the relationship observed previously between biguanide hydrophobicity and inhibitory potency (*12*). Other neutral and highly hydrophobic ligands (DDM, cholate, rotenone, and IACS-2858) have also been observed binding

in overlapping sites in various protein states (fig. S28) (*21*, *22*, *45*). Yet biguanides remain distinctive inhibitors in their state selectivity, shown here and previously (*16*). Of the other inhibitors, only rotenone has been demonstrated to bind to a deactive-like open state, but the binding was not state specific (*21*) and the protein interactions and in vivo effects of biguanides and rotenone are different (*46*, *47*).

Despite the locations of the biguanide- and ubiquinone-binding sites overlapping, metformin does not display classical competitive behavior (*12*). This is likely because neutral, hydrophobic inhibitors such as piericidin A, IACS-2858, and acetogenin (*45*, *48*, *49*) compete for (active-like) enzyme states capable of forming the Michaelis complex, but biguanide binding occurs most readily in deactive-like states that are not preorganized, owing to disordering of the Q-channel loops in NDUFS2, ND3, and ND1 (*24*). The deactive state of complex I is stabilized by high pH (*50*), and we see a strong correlation between pH and inhibitory potency. Weak biguanide inhibition of active complex I could originate either from binding to the Q-channel or from an inhibitory nature to the ND5 lateral helix interaction. The conformation of the NDUFS7 β1–β2 loop that contains Arg^{77} differs between the active, deactive, and slack states. Biguanides bind 7 to 14 Å from Arg^{77} in deactive and slack states, but the arginine guanidinium position is shifted relative to the positions in the DDM- and cholate-bound bovine nanodisc deactive and slack states (PDB IDs 7QSM and 7QSO) (fig. S28) (*22*). In all active models, Arg^{77} is much closer to the putative biguanide-binding site (Fig. 2B), presenting a source of steric hindrance and charge repulsion that acts against strong inhibitory binding and explains the preference of biguanide inhibitors to bind to NEM-sensitive (deactive) complex I [manifesting as a lower IC_{50} for deactivated complex I (Fig. 1B) (*16*)]. Biguanides are less powerful inhibitors when added during catalysis rather than before (*12*), and metformin has been proposed to act by slowing down activation (*51*) rather than by classical inhibition. Our data support an inhibitory mode that primarily acts by preventing reactivation of the resting deactive state, and we conclude that the major inhibitory interaction site of biguanides is the Q-channel.

Local chaotropic drug-protein interactions

A key feature of biguanide interactions with complex I is the displacement and disordering of a portion of the ND5 lateral helix. This type of localized disruption has not been observed previously but may be facilitated by nonspecific interactions with phospholipids (*42*). Considering the similarity of biguanides to the well-known chaotrope and protein denaturant guanidinium, the biguanide may be attracted to the negatively charged ND5-Asp^{554} and ND5-Glu^{559} and specifically destabilize the hydrogen-bonding networks between the lateral helix and subunit ND4. In particular, biguanide binding could weaken the interactions of the lateral helix π-bulge segment, allowing secondary structure shifts to form an α helix and a disordered loop. After this process, the unraveled segment of the helix would no longer provide lateral support to keep the strained junction between ND2 and ND4 (*52*) tightly together, explaining the slight "opening" of the proximal and distal membrane domain interface (fig. S25). Protein stability assays (fig. S2) suggest that complex I is stabilized close to the IM1092 IC_{50} ranges but is destabilized at higher concentrations, and we interpret this behavior to indicate that binding to the Q-channel (and/or ND2-NDUFB5 interface) stabilizes the protein and that chaotropic action occurs at higher concentrations. Our structures in this study are in a detergent micelle with tightly bound lipids present, and the chaotropic interaction site we observe is within the phospholipid headgroup plane. Notably, any inhibitory consequences of structural disturbance in this study are fully reversible, which is demonstrated by full recovery of activity after inhibitor dilution (fig. S2F). We propose that biguanides such as metformin, phenformin, and related drug leads could exert similar local chaotropic actions on any number of cellular proteins, especially membrane proteins, to inhibit or stimulate their usual functions, presenting a novel mode of enzyme-drug interaction and a potential explanation for the breadth of biguanide targets identified thus far.

Implications for in vivo mechanism of action

Although other complex I inhibitor classes have been proposed as possible therapeutic compounds (*53*–*56*), biguanides appear to offer a lower toxicity profile than neutral species such as rotenone, with reduced risk of Parkinsonism (*47*). The key factor may lie in self-limitation of action in the form of negative feedback at the organelle level, meaning that their mechanism of action lowers the risk of complete inhibition of the respiratory chain. The membrane potential leads to increased concentration of biguanides at the complex I site of action, but as complex I is inhibited, the membrane potential falls (*57*), creating an equilibrium that limits the concentration and therefore the degree of inhibition. Toxicity may also be influenced by the ready reversibility of biguanide inhibition seen here and previously (*12*), compared with the irreversibility of very hydrophobic, neutral compounds. In vivo, the active-like state that binds canonical inhibitors is expected to be prevalent in normal tissues (*44*), where the opportunities for biguanides to inhibit may be restricted to the weak binding observed in this study against active-like states, or to the binding of putative catalytic intermediates where NDUFS7-Arg^{77} has moved away from the inhibitory binding site. The deactive state of complex I forms during oxygen starvation (*51*), such as during ischemia (*2*) and in solid tumor microenvironments (*58*). Discrimination of biguanides for this state means that, unlike canonical inhibitors, hypoxic tissues can be selectively targeted with less risk of also compromising mitochondrial respiration in tissues with normally functioning (active) complex I. In such low-oxygen environments, only complex I in the active state may work in the reverse direction, using the PMF and ubiquinol to drive NAD^+ reduction and reactive oxygen species production (*51*). By targeting the deactive state formed in these environments, preventing reactivation, biguanides may diminish reactive oxygen species production by reverse electron transfer, as has been observed previously for metformin (*59*).

Metformin has proven to be a safe and effective biguanide for diabetes treatment, and our data now provide a structure-based rationale for its mode of action on complex I. This work also offers a basis for future structure-based biguanide drug design for the state-specific inhibition of complex I for other potential therapeutic applications (such as cancer treatment) in which metformin has been less successful, as well as improved prediction of additional protein target binding sites for biguanide drugs.

REFERENCES AND NOTES

1. M. Foretz, B. Guigas, L. Bertrand, M. Pollak, B. Viollet, *Cell Metab.* **20**, 953–966 (2014).
2. K. Skemiene *et al.*, *Biomolecules* **10**, 1400 (2020).
3. X. Wang *et al.*, *PLOS ONE* **12**, e0182777 (2017).
4. N. Sato *et al.*, *Respir. Res.* **17**, 107 (2016).
5. S. Lehrer, *World Acad. Sci. J.* **2**, 1 (2020).
6. H. Zhao, K. D. Swanson, B. Zheng, *Trends Cancer* **7**, 714–730 (2021).
7. M. Pollak, *Nat. Med.* **20**, 591–593 (2014).
8. S. Kordes *et al.*, *Lancet Oncol.* **16**, 839–847 (2015).
9. P. J. Goodwin *et al.*, *J. Clin. Oncol.* **31**, 1033–1033 (2013).
10. A. K. Madiraju *et al.*, *Nature* **510**, 542–546 (2014).
11. T. Ma *et al.*, *Nature* **603**, 159–165 (2022).
12. H. R. Bridges, A. J. Y. Jones, M. N. Pollak, J. Hirst, *Biochem. J.* **462**, 475–487 (2014).
13. T. E. LaMoia *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2122287119 (2022).
14. F. Barbieri *et al.*, *Front. Oncol.* **9**, 135 (2019).
15. M. Y. El-Mir *et al.*, *J. Biol. Chem.* **275**, 223–228 (2000).
16. S. Matsuzaki, K. M. Humphries, *Biochemistry* **54**, 2011–2021 (2015).
17. M. R. Owen, E. Doran, A. P. Halestrap, *Biochem. J.* **348**, 607–614 (2000).
18. G. Rena, D. G. Hardie, E. R. Pearson, *Diabetologia* **60**, 1577–1585 (2017).
19. R. A. Miller *et al.*, *Nature* **494**, 256–260 (2013).
20. J. Hirst, *Annu. Rev. Biochem.* **82**, 551–575 (2013).
21. D. Kampjut, L. A. Sazanov, *Science* **370**, eabc4209 (2020).
22. I. Chung *et al.*, *Nat. Commun.* **13**, 2758 (2022).
23. A. A. Agip *et al.*, *Nat. Struct. Mol. Biol.* **25**, 548–556 (2018).
24. J. N. Blaza, K. R. Vinothkumar, J. Hirst, *Structure* **26**, 312–319. e3 (2018).
25. J. Zhu, K. R. Vinothkumar, J. Hirst, *Nature* **536**, 354–358 (2016).
26. T. Sakai *et al.*, *Sci. Rep.* **11**, 4852 (2021).
27. S. W. Kim *et al.*, "Guanidine Compounds and use thereof," US patent application no. 2017/0073331 A1 (16 March 2017).
28. H. R. Bridges, V. A. Sirviö, A. N. A. Agip, J. Hirst, *BMC Biol.* **14**, 65 (2016).

29. M. Babot, A. Birch, P. Labarbuta, A. Galkin, *Biochim. Biophys. Acta* **1837**, 1083–1092 (2014).
30. E. V. Gavrikova, A. D. Vinogradov, *FEBS Lett.* **455**, 36–40 (1999).
31. D. Kathuria, A. A. Bankar, P. V. Bharatam, *J. Mol. Struct.* **1152**, 61–78 (2018).
32. P. V. Afonine *et al*., *Acta Crystallogr. D Struct. Biol.* **74**, 814–840 (2018).
33. J. G. Fedor, A. J. Y. Jones, A. Di Luca, V. R. I. Kaila, J. Hirst, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 12737–12742 (2017).
34. M. J. Landrum *et al*., *Nucleic Acids Res.* **46**, D1062–D1067 (2018).
35. E. L. Blakely *et al*., *Eur. J. Hum. Genet.* **13**, 623–627 (2005).
36. P. Silva-Pinheiro, M. Minczuk, *Nat. Rev. Genet.* **23**, 199–214 (2022).
37. R. G. Efremov, R. Baradaran, L. A. Sazanov, *Nature* **465**, 441–445 (2010).
38. J. Ludwiczak *et al*., *Sci. Rep.* **9**, 6888 (2019).
39. J. P. Cartailler, H. Luecke, *Structure* **12**, 133–144 (2004).
40. F. Davidoff, *J. Biol. Chem.* **246**, 4017–4027 (1971).
41. M. J. MacDonald, I. H. Ansari, M. J. Longacre, S. W. Stoker, *Diabetes* **70**, 1575–1580 (2021).
42. G. Schäfer, *Biochem. Pharmacol.* **25**, 2005–2014 (1976).
43. N. Samart, C. N. Beuning, K. J. Haller, C. D. Rithner, D. C. Crans, *Langmuir* **30**, 8697–8706 (2014).
44. Z. Yin *et al*., *Nat. Commun.* **12**, 707 (2021).
45. I. Chung *et al*., *Sci. Adv.* **7**, eabg4000 (2021).
46. M. Konopleva *et al*., Targeting Oxidative Phosphorylation with a Mitochondrial Complex I Inhibitor is limited by Mechanism-based Toxicity, 1506700/v1, Research Square [Preprint] (2022).
47. S. Heinz *et al*., *Arch. Toxicol.* **93**, 2603–2615 (2019).
48. H. R. Bridges *et al*., *Nat. Commun.* **11**, 5261 (2020).
49. D. N. Grba *et al*., *J. Biol. Chem.* **298**, 101602 (2022).
50. A. B. Kotlyar, V. D. Sled, A. D. Vinogradov, *Biochim. Biophys. Acta* **1098**, 144–150 (1992).
51. A. Galkin, S. Moncada, *Interface Focus* **7**, 20160104 (2017).
52. A. Di Luca, V. R. I. Kaila, *Biochim. Biophys. Acta Bioenerg.* **1862**, 148382 (2021).
53. N. Madhusudhan *et al*., *ACS Chem. Biol.* **15**, 158–170 (2020).
54. J. R. Molina *et al*., *Nat. Med.* **24**, 1036–1046 (2018).
55. A. Naguib *et al*., *Cell Rep.* **23**, 58–67 (2018).
56. D. A. Russell *et al*., *J. Nat. Prod.* **83**, 1829–1845 (2020).
57. Z. Drahota *et al*., *Physiol. Res.* **63**, 1–11 (2014).
58. A. Emami Nejad *et al*., *Cancer Cell Int.* **21**, 62 (2021).
59. C. Batandier *et al*., *J. Bioenerg. Biomembr.* **38**, 33–42 (2006).

ACKNOWLEDGMENTS

We thank D. Chirgadze (Cambridge), J. Radecke, and Z. Yang (eBIC) for their assistance with grid screening and data collection; Diamond for access and support of the cryo-EM facilities at the UK national electron Bio-Imaging Centre (eBIC), proposals BI22238-10 and EM17057-23, funded by the Wellcome Trust, MRC, and BBSRC; members of the Hirst laboratory, especially N. Agip, for their participation in bovine mitochondrial preparations; S. McLaughlin (LMB, Cambridge) for access to the nano-DSF; T. Terwilliger (Los Alamos, NM) for guidance with map sharpening and composite map generation; S. Ding (MBU) for recording mass spectrometry data; H. Prag (MBU) for recording liquid chromatography data; P. Gierth (Cambridge) for recording nuclear magnetic resonance data; J. Bosak (MBU) for culturing and authenticating cells; and M. Iadanza (Leeds) and P. Afanasyev (ETH, Zurich) for their scripts deposited on Github. **Funding:** This work was supported by the Medical Research Council (MC_U105663141 and MC_UU_00015/2 to J.H.) and by ImmunoMet Therapeutics Inc. **Author contributions:** Conceptualization: H.R.B., M.N.P., and J.H. Methodology: H.R.B. and J.N.B. Validation: H.R.B. Formal analysis: H.R.B. Investigation: H.R.B., J.N.B., Z.Y., and I.C. Data curation: H.R.B. Writing – original draft: H.R.B. Writing – review and editing: H.R.B., J.N.B., Z.Y., I.C., M.N.P., and J.H. Visualization: H.R.B. Supervision: J.H. Project administration: H.R.B. and J.H. Funding acquisition: J.H. **Competing interests:** ImmunoMet Therapeutics Inc. are the inventors on patent US2017/007331 for the biguanide compound IM1761092 used in this study. M.N.P. is on the scientific advisory board of ImmunoMet Therapeutics and holds shares in the company. The authors declare that they have no further competing interests. **Data and materials availability:** The structure data accession codes are EMD-14251, EMD-14252, EMD-14253, EMD-15254, EMD-15355, and PDB-7R41 (Active-1092-i); EMD-4256, EMD-14257, EMD-14258, EMD-14259, EMD-14260, and PDB-7R42 (Active-1092-ii); EMD-14261, EMD-14262, EMD-14263, EMD-14264, EMD-14265, and PDB-7R43 (Active-1092-iii); EMD-14266, EMD14267, EMD-14268, EMD-14269, EMD-14270, and PDB-7R44 (Active-1092-iv); EMD-4272, EMD-14273, EMD-14274, EMD14275, EMD14276, and PDB-7R45 (Deactive-1092-i); EMD-14277, EMD-14278, EMD14279, EMD-14280, EMD-14281, and PDB-7R46 (Deactive-1092-ii); EMD-14282, EMD-14283, EMD-14284, EMD-14285, EMD-14286, and PDB-7R47 (Deactive-1092-iii); EMD-14287, EMD-14288, EMD-14289, EMD-14290, EMD-14291, and PDB-7R48 (Deactive-1092-iv); EMDB-14292, EMD-14293, EMD-14294, EMD-14295, EMD-14296, and PDB-7R4C (Deactive-1092-v); EMDB-14297, EMD-14298, EMD-14299, EMD-14300, EMD-14301, and PDB-7R4D (Deactive-1092-vi); EMDB-14302, EMD-14304, EMD-14305, EMD-14306, and PDB-7R4F (Slack-1092-i); EMD-14307, EMD-14308, EMD-14309, EMD-14310, EMD-14311, and PDB-7R4G (Slack-1092-ii); EMBD-14127, EMD-14128, EMD-14129, EMD14130, EMD-14131, and PDB-7QSD (Inhibitor-free active); and EMDB-14126 (Inhibitor-free slack). Raw micrograph images are available at EMPIAR with accession codes EMPIAR-10991 (in presence of IM1761092) and EMPIAR-10984 (inhibitor-free). Otherwise, all data needed to evaluate the conclusions in the paper are present in the paper and/or the supplementary materials. IM1092 was supplied by ImmunoMet Therapeutics Inc. by collaboration agreement. ImmunoMet Therapeutics Inc. accepts proposals to supply IM1092 and ~1000 other biguanide compounds from their biguanide library for research purposes. Such proposals should be directed to dwelsch@immunomet.com.

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade3332
Materials and Methods
Supplementary Text
Figs. S1 to S28
Tables S1 to S15
References (*60*–*78*)
MDAR Reproducibility Checklist

Submitted 10 August 2022; accepted 23 December 2022
10.1126/science.ade3332

CRYSTALLINE ORDERING

Truchet-tile structure of a topologically aperiodic metal–organic framework

Emily G. Meekel[1], Ella M. Schmidt[1,2], Lisa J. Cameron[3], A. David Dharma[3], Hunter J. Windsor[3], Samuel G. Duyker[3,4], Arianna Minelli[1], Tom Pope[5], Giovanni Orazio Lepore[6], Ben Slater[5], Cameron J. Kepert[3]*, Andrew L. Goodwin[1]*

When tiles decorated to lower their symmetry are joined together, they can form aperiodic and labyrinthine patterns. Such Truchet tilings offer an efficient mechanism of visual data storage related to that used in barcodes and QR codes. We show that the crystalline metal–organic framework $[OZn_4][1,3\text{-benzenedicarboxylate}]_3$ (TRUMOF-1) is an atomic-scale realization of a complex three-dimensional Truchet tiling. Its crystal structure consists of a periodically arranged assembly of identical zinc-containing clusters connected uniformly in a well-defined but disordered fashion to give a topologically aperiodic microporous network. We suggest that this unusual structure emerges as a consequence of geometric frustration in the chemical building units from which it is assembled.

In 1704, Sébastian Truchet described a variety of visually appealing patterns generated by a single square tile painted differently on either side of its diagonal (*1*). Truchet always arranged his tiles on the same square lattice but obtained different patterns by varying the orientations of each tile. Truchet tilings have been generalized to encompass any periodic covering of space (partial or complete) with one or more tiles that have been decorated to reduce their symmetry (*2*). Breaking color-inversion symmetry, for example, gives barcodes and QR codes as simple Truchet tilings in one and two dimensions, respectively. Each strip or pixel encodes a binary state that is exploited in their application as machine-readable information stores (Fig. 1A). More-complex Truchet tilings store information in aperiodic patterns of loops and networks in two or three dimensions (Fig. 1, B and C) (*3*).

Because the solid-state structures of many crystalline materials can be related to periodic tilings of three-dimensional (3D) space (*4*), we explored the possibility of realizing an atomic-scale Truchet tiling in metal–organic frameworks (MOFs), the structures of which can be controlled through judicious selection of the nodes and linkers from which they are assembled (*5*). In the canonical system MOF-5 $[OZn_4][1,4\text{-bdc}]_3$, the combination of OZn_4 as the octahedrally coordinating node and benzenedicarboxylate (bdc) as the connecting linker generates a structure with the simple cubic topology (*6*). This structure can be mapped to the tiling of space by cubes (*7*).

Synthesis and characterization of TRUMOF-1

Greater complexity can be introduced into MOFs by lowering the symmetry of their building blocks (*8*, *9*), and we synthesized a new derivative of MOF-5 in which the linear

[1]Inorganic Chemistry Laboratory, University of Oxford, Oxford OX1 3QR, UK. [2]Fachbereich Geowissenschaften, Universität Bremen, D-28359 Bremen, Germany. [3]School of Chemistry, University of Sydney, New South Wales 2006, Australia. [4]Sydney Analytical, Core Research Facilities, University of Sydney, New South Wales 2006, Australia. [5]Department of Chemistry, University College London, London WC1H 0AJ, UK. [6]Department of Earth Sciences, University of Florence, 50121 Florence, Italy.
***Corresponding author. Email: cameron.kepert@sydney.edu.au (C.J.K.); andrew.goodwin@chem.ox.ac.uk (A.L.G.)**

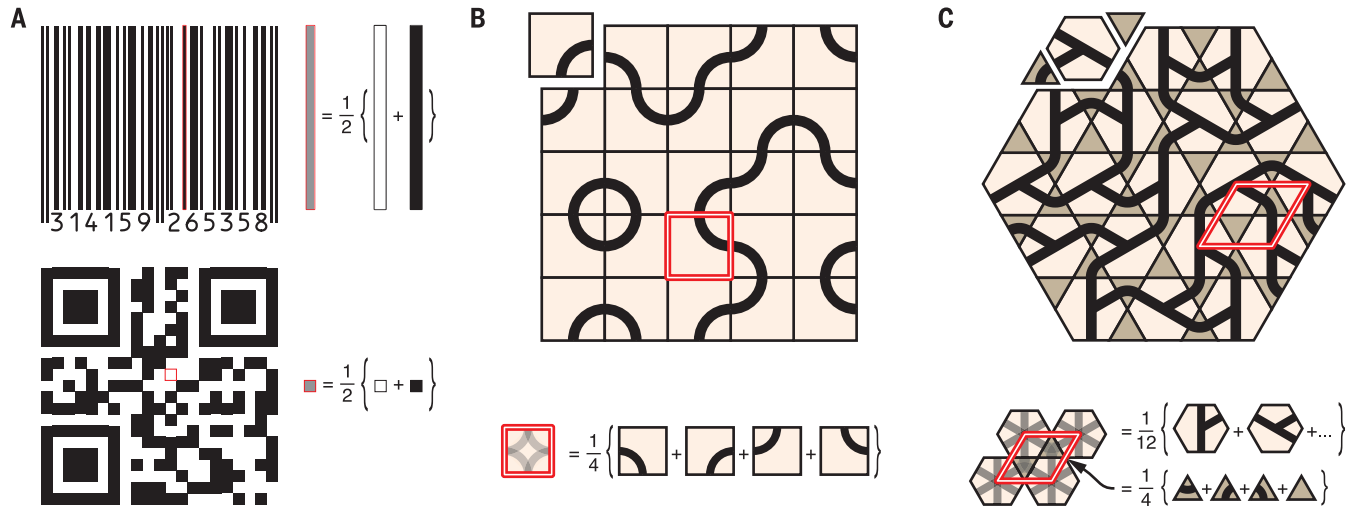


Fig. 1. Representative Truchet tilings. (**A**) Barcodes and QR codes store information in tilings of strips or pixels that break color-inversion symmetry. The average tile color, gray, is invariant under color inversion, but individual strips or pixels (examples outlined in red) store information in the way in which color-inversion symmetry is broken. The patterns formed are Truchet tilings in the sense that individual tiles are lower in symmetry than the configurational average. (**B**) Decorating square tiles with a single arc can generate Truchet tilings with curved paths. The positions of the square tiles are translationally periodic, but the continuous pattern produced by the arc decorations is aperiodic. Projecting any such tiling onto the underlying unit cell (shown in red) gives an average structure with partial occupancies. For large systems, one expects equal occupancies of each of the four possible arc orientations. (**C**) More-complex Truchet tilings can be generated using a combination of different decorated and undecorated tiles. In this particular example, based on the trihexagonal tiling of the plane, the connected path of line decorations forms a single percolating network that spans the entire system.

1,4-bdc linker was replaced by its bent 1,3-bdc analog. The crystalline material that we obtained, which we call TRUMOF-1, produced an x-ray diffraction (XRD) pattern that included both conventional Bragg diffraction reflections and very weak diffuse scattering. Because the underlying chemistries of TRUMOF-1 and MOF-5 were so similar, we anticipated that the composition of TRUMOF-1 should be $[OZn_4][1,3\text{-bdc}]_3$. Structure solution from the Bragg component (sensitive only to the configurational average) gave the result shown in Fig. 2A: $[OZn_4]$ clusters were positioned at the corners and face centers of the $F\bar{4}3m$ cubic unit cell and connected by what appeared to be trigonally symmetric linker molecules that were only partly occupied. Distinguishable in the Fourier maps at the linker sites are benzene C atoms and carboxylate C/O atoms, the occupancies of which we refined to values of 0.70(4) and 0.46(3), respectively. These occupancies implied a crystal composition of $[OZn_4][1,3\text{-bdc}]_{2.8(2)}$ and are near the values (0.75 and 0.5) expected for the idealized composition $[OZn_4][1,3\text{-bdc}]_3$. The configurational averages of Truchet tilings necessarily contain features with partial occupancies (Fig. 1).

We were able to synthesize TRUMOF-1 in bulk, and powder samples produced XRD patterns entirely consistent with our single-crystal measurements (fig. S6). The framework pores of as-synthesized samples contained solvent that could be exchanged with other solvents or removed by heating under vacuum. The empty framework adsorbed a range of gases (including CO_2) and had a Brunauer-Emmett-Teller (BET) surface area of 765 $m^2\ g^{-1}$ (fig. S12). Its thermal stability was similar to that of MOF-5, but its adsorption enthalpies were about twice as large (figs. S4 and S13), which is indicative of the presence of smaller pores. In all of these various respects, TRUMOF-1 behaves as a conventional MOF (*10*).

We interpreted the average crystal structure of TRUMOF-1 by first considering the partially occupied ligand site. ^1H and ^{13}C nuclear magnetic resonance spectra of acid-digested TRUMOF-1 samples revealed the only bridging ligand to be 1,3-bdc (fig. S2), so the crystallographic model must represent an average over three equally populated orientations of the 1,3-bdc linker (Fig. 2B). The emergence of an apparently threefold-symmetric ligand resulted from the crystallographic site having higher point symmetry than the linker itself. To give the chemically sensible stoichiometry $[OZn_4][1,3\text{-bdc}]_3$, the site must also be vacant with 25% probability; note that this value was consistent with the observed site occupancies.

By contrast, the OZn_4 cluster appeared ordered and fully occupied. The cluster was surrounded by 12 carboxylate anion sites, each occupied with 50% probability. Zinc K-edge extended x-ray fine structure (EXAFS) measurements showed the Zn coordination environment to be essentially identical to that in the octahedrally coordinated OZn_4 clusters of MOF-5 (*11*) (fig. S11). We incorporated this finding into the crystallographic model in terms of a configurational average over 24 symmetry-related octahedral-coordination decorations of the cluster (Fig. 2C). Again, each cluster decoration had lower point symmetry than the crystallographic site on which the cluster sat (the carboxylate positions broke ideal octahedral symmetry; fig. S15). On purely geometric grounds, the various partial site occupancies should be strongly correlated, because the orientation of a given cluster decoration placed strong constraints on which neighboring 1,3-bdc linkers were present and in which possible orientations they could lie (*12*, *13*).

Mapping to a Truchet tiling

Both local symmetry-lowering and orientational-matching rules allowed us to map the structure to a 3D Truchet tiling, and we tested this model both computationally and experimentally. To obtain the underlying tiling, we used a Voronoi decomposition of the crystal structure based on the average cluster and ligand positions. The two tile shapes so produced, a cube and a trigonal bipyramid, filled 3D space when placed on the TRUMOF-1 cluster and ligand sites, respectively. The Truchet decorations are given by the symmetry lowering implied by 1,3-bdc occupancies in the TRUMOF-1 crystal structure (Fig. 2D). For example, each cube is decorated by a pseudo-octahedral arrangement of rods that capture the arrangement of carboxylate substituents around each OZn_4 cluster. Likewise, three-quarters of the trigonal bipyramids were decorated by an arc

that reflected the bent linking geometry of a 1,3-bdc molecule. The remaining quarter of the trigonal bipyramidal tiles were left undecorated. Four triangular bipyramids packed together to form a cube. Any combination of orientations of these different tiles, alternating on their respective sites, that produced a connected network then would relate to a chemically sensible model of 1,3-bdc–linked OZn_4 clusters. Each of these clusters was octahedrally coordinated by six 1,3-bdc linkers, and each 1,3-bdc linker connected exactly two clusters (Fig. 2E). There are infinitely many such networks.

To test whether this Truchet-tile description of TRUMOF-1 is physically reasonable, we performed a series of ab initio density functional theory (DFT) calculations. Our approach was to calculate the geometry and cell-optimized 0 K energies of periodic realizations ("approximants") (*14*) of the system with increasingly large unit cells so as to understand the underlying energetic landscape. There were only three symmetry-distinct 1 × 1 × 1 approximants, that is, those with unit cells commensurate with the $(\sim 15\ \text{Å})^3$ cell of the average structure. Each relaxed to give a sensible structure like that shown in Fig. 2E, with 10 distinct realizations giving equilibrium energies within a few kilojoules per mole of Zn of one another.

The number of possible configurations grew very quickly with system size. For the 2 × 2 × 2 case (32 formula units; 1696 atoms), there were too many possibilities to enumerate, so we focused on 20 representative realizations. These larger configurations spanned a wider range of energies and had more open structures (Fig. 3A); the range of unit-cell volumes observed was characteristic of flexible MOF structures (*15*). Because the enthalpic differences among these periodic TRUMOF-1 approximants were commensurate with solvent-binding enthalpies in zinc carboxylate MOFs (such as MOF-5) (*16*, *17*), we expected that guest inclusion would reduce the spread of energies in practice because loading would be higher in the less-dense realizations. Therefore, we concluded that the configurational landscape accessible to TRUMOF-1 spanning different Truchet-tile arrangements involved many thermally accessible minima. This finding helped to explain why synthesis at ~400 K generated a disordered rather than an ordered configuration. Aperiodic networks would also be favored on entropic grounds.

Single-crystal diffuse scattering

The experimental measurement with the greatest sensitivity to structural disorder and its correlation in crystalline materials is single-crystal diffuse scattering (*18*). All disordered materials give rise to diffuse scattering, but the presence of structure in this scattering implies nonrandom disorder, and its form can be diagnostic of specific correlations governing local structure (*19*, *20*). The experimental x-ray diffuse scattering pattern of TRUMOF-1 was dominated by transverse-polarized diagonal features characteristic of correlated displacements (Fig. 3B) (*21*). Because we did not find any strong temperature dependence to the scattering, this contribution could be attributed predominantly to static displacements rather than to thermal effects. The corresponding

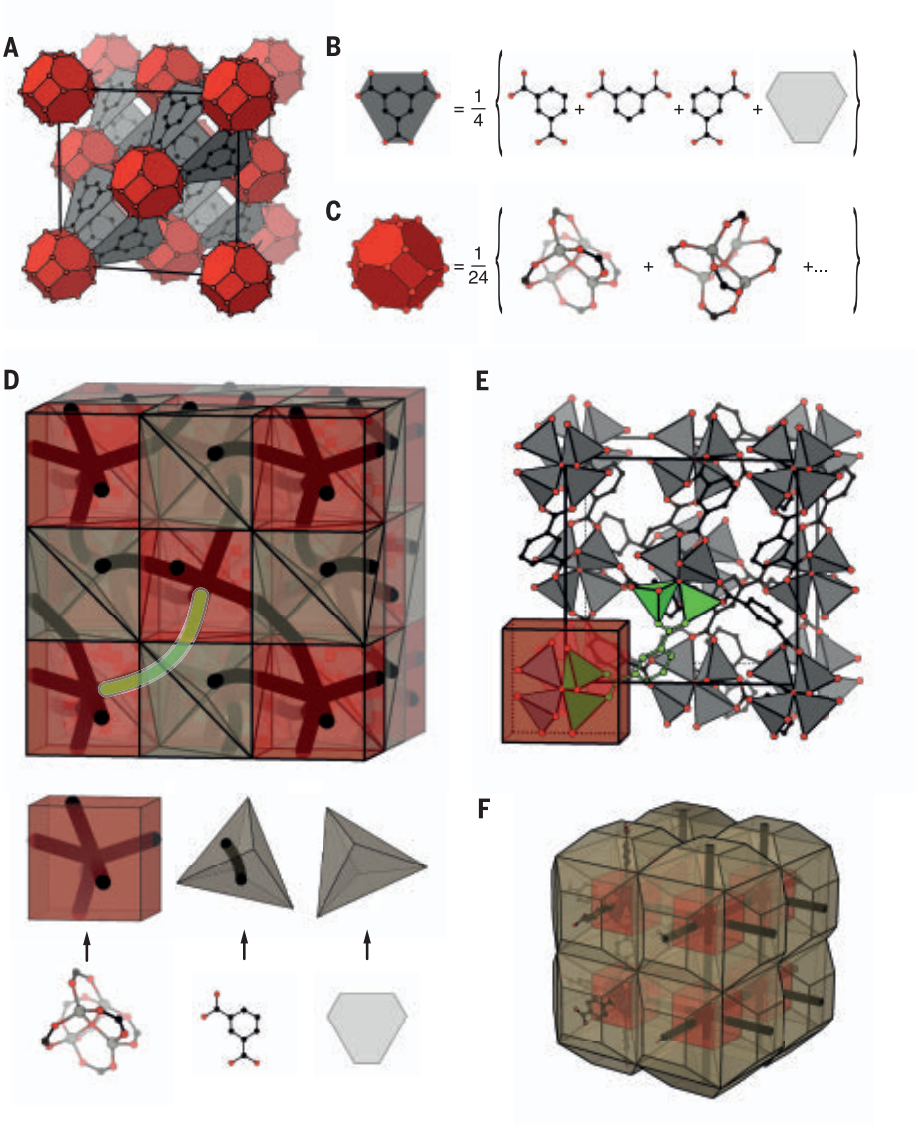


Fig. 2. Crystal structure of TRUMOF-1. (**A**) Representation of the average structure of TRUMOF-1, with OZn_4-centered clusters shown as red cuboctahedra and trigonally symmetric ligand sites shown as gray triangles. (**B**) Interpretation of the crystallographic model at the ligand site in terms of a configurational average over three 1,3-bdc orientations and 25% linker vacancies. C atoms are shown in black, O atoms in red, and H atoms are omitted for clarity. (**C**) Related interpretation of the inorganic cluster as an average over 24 octahedral cluster orientations. Colors are as in (B), with Zn atoms shown in gray. (**D**) The Truchet tiles used in our mapping correspond to (from left to right) the pseudo-octahedral inorganic cluster orientations, the 1,3-bdc linker orientations, and the positions of ligand vacancies. Connecting these tiles to fill 3D space forms networks that, like the patterns shown in Fig. 1, are aperiodic. Note that the trigonal bipyramids pack such that any four combine to form a cube of the same dimensions as the (red) cluster cube. (**E**) One possible 1 × 1 × 1 approximant of the TRUMOF-1 structure with atom coordinates and cell geometry optimized using DFT. Note that the cluster positions are as in (A), but not all linker sites are occupied. The relation between cluster and cubic Truchet tile is emphasized for one cluster; likewise, the connectivity between two clusters in the foreground and the corresponding link in the related Truchet tiling (D) are highlighted in green. (**F**) Representation of the (ordered) structure of MOF-5 in terms of its underlying tiling for comparison against that given for TRUMOF-1 in (D).

3D difference pair-distribution function (3D-ΔPDF) captured the real-space correlations from which the diffuse scattering arose (Fig. 3C) (*22*). The diffuse scattering and 3D-ΔPDF calculated from our larger DFT-relaxed configurations closely matched experimental results (Fig. 3, B and C), indicating that the Truchet-tile model from which these configurations are generated captured the fundamental nature of the correlated disorder in TRUMOF-1. The subtle differences were attributed to the finite size of our DFT model—and thus the sharpness of features in reciprocal space—and the role of dynamical effects, most obviously affecting low-|**r**| features in the 3D-ΔPDF. We anticipate even better agreement for larger approximants (were they computationally accessible), which would more closely approximate the aperiodic limit. A more detailed discussion of the sensitivity of the diffuse scattering to different aspects of our Truchet-tile model is given in the supplementary materials.

As a further check against experimental results, we show in Fig. 3D that the distribution of atomic positions determined from configurational averaging over our DFT configurations was consistent with the anisotropic displacement parameters determined in our single-crystal XRD refinements. Our model indicated that static disorder contributed ~0.1 Å^2 to atomic mean-squared displacements in the $F\bar{4}3m$ average structure model. A similar value was obtained by extrapolating to 0 K the experimental displacement parameters measured between 100 and 350 K (fig. S18). Therefore, our Truchet-tile model rationalizes at once both average and local structure measurements of the structure of TRUMOF-1.

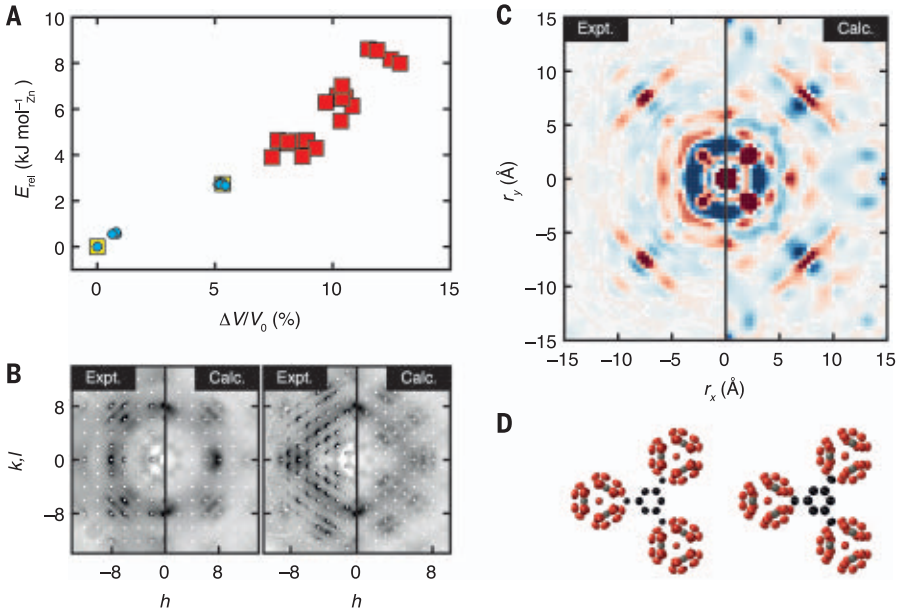


Fig. 3. Computational and experimental validation of the TRUMOF-1 structure. (**A**) The relative lattice energies of different 1 × 1 × 1 (blue circles) and 2 × 2 × 2 (red squares) approximants are within a few kilojoules per mole of one another. Also included are the energies for geometry-optimized 2 × 2 × 2 supercells of two representative 1 × 1 × 1 approximants (yellow squares). The more-disordered configurations are those at higher energy, for which the molar volume is largest. (**B**) Comparison of experimental and calculated single-crystal x-ray diffuse scattering in the (left) (*hk*0) and (right) (*hhl*) planes. The Bragg reflection positions are masked by white circles. (**C**) Comparison between experimental and calculated r_xr_y0 cuts of the 3D-ΔPDF. Positive correlations are shown in red and negative correlations in blue. (**D**) Comparison between the experimentally derived crystallographic model for scattering density at the ligand site at 100 K (left) and that obtained by projection of DFT-relaxed structures onto the same site (right). Ellipsoids are shown at 50% probability.

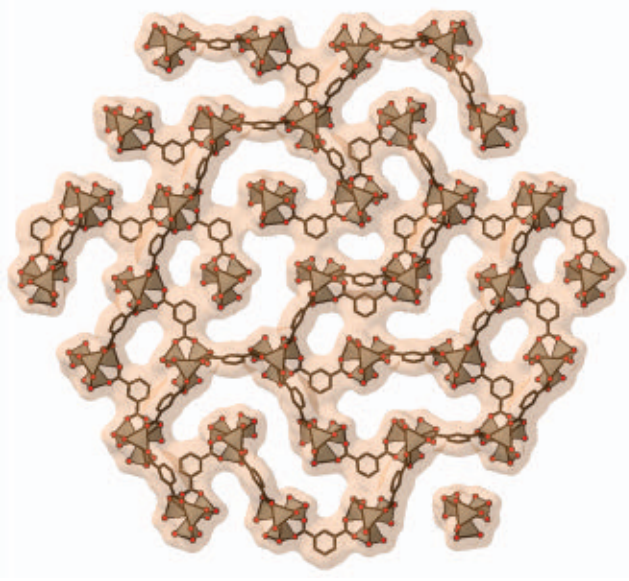


Fig. 4. Disordered pore network structure. A representative planar fragment of a TRUMOF-1 structure model is illustrated viewed down the [111] crystal axis. Only those 1,3-bdc molecules that link clusters within this plane are shown. Each cluster is connected to six neighboring clusters, some of which will be in planes above or below that shown here. The surface of the solvent-accessible pores is shown as a gray outline. Note that, despite the periodic arrangement of inorganic clusters, the pore network is disordered. There is a qualitative similarity to the network structure shown in Fig. 1C.

Discussion

The topology of TRUMOF-1 is unusual in a number of respects. The MOF network is a fully six-connected net with nodes arranged on a periodic lattice (as in MOF-5), but its connectivity is aperiodic (in contrast to MOF-5). No two crystals would have the same network connectivity—that is, they store different information—despite sharing the same statistical properties. The system is intermediate to conventional crystalline MOFs and amorphous MOF derivatives with continuous random network structures (*23*, *24*). The topological aperiodicity of TRUMOF-1 may be important because it translates into a disordered pore network with specific characteristics (*25*). We found that guest molecules were sensitive to this aperiodicity, because loading TRUMOF-1 crystals with I_2 enhanced the x-ray diffuse scattering contrast (fig. S9). A representative pore structure extracted from one of our DFT configurations is illustrated in Fig. 4. The complex transport pathways that we observed were qualitatively similar to those in gels (*26*) and granular porous media (*27*) and thus may be useful in separations. We found good consistency between experimental adsorption behavior and that calculated from grand canonical Monte Carlo simulations based on our structural model; as anticipated, the smaller pore dimensions relative to MOF-5 reflect the stronger adsorption enthalpies of TRUMOF-1 (fig. S14).

There is no reason to expect that TRUMOF-1 should be the only MOF with a Truchet-tile structure. Because every conventional periodic tiling can be modified to give Truchet tilings, suitable symmetry lowering of inorganic and organic building units should allow targeted synthesis of alternative Truchet MOFs. Indeed, the structures of some previously described MOFs (*28*) and related systems (*29*) with partially occupied and/or complex structures might be reinterpretable as Truchet tilings. We note that TRUMOF-1 was assembled from an inorganic component predisposed to forming extended network structures (*30*) and an organic component usually used in the formation of discrete molecular cages (*31*). The concept of geometrically frustrating the competing structure-directing effects of two synthons to drive a complex state may generalize beyond this particular system (*32*). The discovery of a crystalline Truchet-tile material allows experimental investigation of the impact of aperiodic connectivity on collective materials properties. Aperiodicity could frustrate elastic instabilities (*33*) or dampen long-wavelength acoustic phonons, as occurs in superlattice heterostructures (*34*) and as required for optimizing thermoelectric response (*35*). Therefore, Truchet-tile–inspired systems may provide a useful avenue for the design of a variety of functional systems.

REFERENCES AND NOTES

1. S. Truchet, *Mém. Acad. R. Sci. (Paris)* **1704**, 363–372 (1704).
2. C. S. Smith, P. Boucher, *Leonardo* **20**, 373–385 (1987).
3. C. Browne, *Comput. Graph.* **32**, 268–281 (2008).
4. V. A. Blatov, O. Delgado-Friedrichs, M. O'Keeffe, D. M. Proserpio, *Acta Crystallogr. A* **63**, 418–425 (2007).
5. O. Delgado-Friedrichs, M. O'Keeffe, O. M. Yaghi, *Phys. Chem. Chem. Phys.* **9**, 1035–1043 (2007).
6. H. Li, M. Eddaoudi, M. O'Keeffe, O. M. Yaghi, *Nature* **402**, 276–279 (1999).
7. M. O'Keeffe, *APL Mater.* **2**, 124106 (2014).
8. V. Guillerm, D. Maspoch, *J. Am. Chem. Soc.* **141**, 16517–16538 (2019).
9. E. G. Meekel, A. L. Goodwin, *CrystEngComm* **23**, 2915–2922 (2021).
10. M. Eddaoudi, H. Li, O. M. Yaghi, *J. Am. Chem. Soc.* **122**, 1391–1397 (2000).
11. J. Hafizovic *et al.*, *J. Am. Chem. Soc.* **129**, 3612–3620 (2007).
12. A. Simonov, A. L. Goodwin, *Nat. Rev. Chem.* **4**, 657–673 (2020).
13. S. Ehrling *et al.*, *Nat. Chem.* **13**, 568–574 (2021).
14. P. M. M. Thygesen *et al.*, *Phys. Rev. Lett.* **118**, 067201 (2017).
15. S. Krause *et al.*, *Nature* **532**, 348–352 (2016).
16. J. T. Hughes, A. Navrotsky, *J. Am. Chem. Soc.* **133**, 9184–9187 (2011).
17. C. K. Brozek *et al.*, *ACS Cent. Sci.* **1**, 252–260 (2015).
18. T. R. Welberry, *Rep. Prog. Phys.* **48**, 1543–1594 (1985).
19. D. A. Keen, A. L. Goodwin, *Nature* **521**, 303–309 (2015).
20. A. Simonov *et al.*, *Nature* **578**, 256–260 (2020).
21. T. R. Welberry, T. Weber, *Crystallogr. Rev.* **22**, 2–78 (2016).
22. T. Weber, A. Simonov, *Z. Kristallogr.* **227**, 238–247 (2012).
23. A. R. Overy *et al.*, *Nat. Commun.* **7**, 10445 (2016).
24. T. D. Bennett *et al.*, *Phys. Rev. Lett.* **104**, 115503 (2010).
25. D. Ormrod Morley, A. L. Goodwin, M. Wilson, *Phys. Rev. E* **102**, 062308 (2020).
26. X. Li, T. Sakai, in *Physics of Polymer Gels*, T. Sakai, Ed. (Wiley, 2020); pp. 137–150.
27. F. A. L. Dullien, *Porous Media: Fluid Transport and Pore Structure* (Academic, ed. 2, 1992).
28. S. Yang *et al.*, *Nat. Mater.* **11**, 710–716 (2012).
29. M. O. Blunt *et al.*, *Science* **322**, 1077–1081 (2008).
30. M. Eddaoudi *et al.*, *Science* **295**, 469–472 (2002).
31. M. Eddaoudi *et al.*, *J. Am. Chem. Soc.* **123**, 4368–4369 (2001).
32. X. Zeng *et al.*, *Science* **331**, 1302–1306 (2011).
33. M. Moshe, M. J. Bowick, M. C. Marchetti, *Phys. Rev. Lett.* **120**, 268105 (2018).
34. R. Venkatasubramanian, E. Siivola, T. Colpitts, B. O'Quinn, *Nature* **413**, 597–602 (2001).
35. Q. D. Gibson *et al.*, *Science* **373**, 1017–1022 (2021).

ACKNOWLEDGMENTS

We acknowledge the provision of beamtime by the LISA beamline at the ESRF and the facilities and the scientific and technical assistance of Sydney Analytical, a core research facility at the University of Sydney. **Funding:** This work was supported by the Australian Research Council (grants DP150104620 and DP190103130 to C.J.K.), the Royal Society (industry fellowship grant IF160062 to B.S. and studentship INF\PHD\180018 to T.P.), and the European Research Council (advanced grant 788144 to A.L.G.). The authors also acknowledge support in the form of access to the UK's ARCHER2 supercomputer through the Materials Chemistry Consortium, which is funded by the Engineering and Physical Sciences Research Council (grant EP/R029431). **Author contributions:** E.G.M. and L.J.C. synthesized and characterized the materials. E.G.M., E.M.S., and A.M. measured the x-ray diffuse scattering data, which were then interpreted by E.G.M., E.M.S., A.M., and A.L.G. E.G.M., E.M.S., and A.L.G. developed the Truchet-tile model. E.G.M., L.J.C., A.D.D., H.J.W., S.G.D., C.J.K. and A.L.G. performed the crystallographic measurements and analysis. E.G.M., L.J.C., A.D.D. and C.J.K. performed and analyzed the gas adsorption measurements. G.O.L. performed the EXAFS measurements, which were interpreted by E.G.M., G.O.L., and A.L.G. T.P. and B.S. performed and interpreted the DFT and GCMC calculations. C.J.K. and A.L.G. conceived and coordinated the project. A.L.G. wrote the manuscript with input from all authors. All authors participated in discussions regarding the science presented in the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All crystallographic data have been deposited with the CCDC, with reference numbers 2203467 to 2203473. All other data needed to evaluate the conclusions in the study are present in the main manuscript or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade5239
Materials and Methods
Supplementary Text
Figs. S1 to S18
Tables S1 to S4
References (*36*–*62*)
Data Files S1 to S6

Submitted 26 August 2022; accepted 17 November 2022
10.1126/science.ade5239

BIOSYNTHESIS

Complex scaffold remodeling in plant triterpene biosynthesis

Ricardo De La Peña[1]†, Hannah Hodgson[2]†, Jack Chun-Ting Liu[3]†, Michael J. Stephenson[4], Azahara C. Martin[5], Charlotte Owen[2], Alex Harkess[6], Jim Leebens-Mack[7], Luis E. Jimenez[1], Anne Osbourn[2]*, Elizabeth S. Sattely[1,8]*

Triterpenes with complex scaffold modifications are widespread in the plant kingdom. Limonoids are an exemplary family that are responsible for the bitter taste in citrus (e.g., limonin) and the active constituents of neem oil, a widely used bioinsecticide (e.g., azadirachtin). Despite the commercial value of limonoids, a complete biosynthetic route has not been described. We report the discovery of 22 enzymes, including a pair of neofunctionalized sterol isomerases, that catalyze 12 distinct reactions in the total biosynthesis of kihadalactone A and azadirone, products that bear the signature limonoid furan. These results enable access to valuable limonoids and provide a template for discovery and reconstitution of triterpene biosynthetic pathways in plants that require multiple skeletal rearrangements and oxidations.

Among numerous complex triterpenes that are found in the plant kingdom, limonoids are particularly notable given their wide range of biological activities and structural diversity that stems from extensive scaffold modifications (*1*, *2*). Produced by mainly two families in the Sapindales, Rutaceae (citrus) and Meliaceae (mahogany) (*3*), these molecules bear a signature furan and include more than 2800 known structures (*4*, *5*). Azadirachtin, a well-studied limonoid, exemplifies the substantial synthetic challenge for this group of molecules, with 16 stereocenters and 7 quaternary carbons. Few synthetic routes to limonoids have been reported (*6*–*8*), and, more generally, complete biosynthetic pathways to triterpenes with extensive scaffold modifications have remained

[1]Department of Chemical Engineering, Stanford University, Stanford, CA 94305, USA. [2]Department of Biochemistry and Metabolism, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK. [3]Department of Chemistry, Stanford University, Stanford, CA 94305, USA. [4]School of Chemistry, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK. [5]Department of Crop Genetics, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK. [6]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA. [7]Department of Plant Biology, 4505 Miller Plant Sciences, University of Georgia, Athens, GA 30602, USA. [8]Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305, USA.
***Corresponding author. Email: anne.osbourn@jic.ac.uk (A.O.); sattely@stanford.edu (E.S.S.)**
†These authors contributed equally to this work.

elusive. This lack of production routes limits the utility and biological investigation of clinical candidates from this diverse compound class (*9*).

Around 90 limonoids have also been reported to have anti-insect activity (*2*), and several have also been found to target mammalian receptors and pathways (*4*). For example, azadirachtin (Fig. 1), the main component of biopesticides derived from the neem tree (*Azadirachta indica*), is a potent antifeedant, active against >600 insect species (*9*). Perhaps related to antifeedant activity, Rutaceae limonoids such as nomilin, obacunone, and limonin (Fig. 1) that accumulate in *Citrus* species at high levels (*3*) are partially responsible for the "delayed bitterness" of citrus fruit juice, which causes serious economic losses for the citrus juice industry worldwide (*10*). In mammalian systems, several limonoids have shown inhibition of HIV-1 replication (*11*) and anti-inflammatory activity (*12*). Some limonoids of pharmaceutical interest have also been associated with specific mechanisms of action: Gedunin (Fig. 1) and nimbolide (fig. S1) exert potent anticancer activity through Hsp90 inhibition (*13*) and RNF114 blockade (*14*, *15*), respectively.

Limonoids are unusual within the triterpene class because of their extensive biosynthetic scaffold rearrangements. They are referred to as tetranortriterpenoids because their signature tetracyclic, triterpene scaffold (protolimonoid), loses four carbons during the formation of a signature furan ring to yield the basic C-26 limonoid structure (Fig. 1). A range of modifications can then occur to the basic limonoid scaffold through the cleavage of one or more of the four main rings (fig. S1) (*16*, *17*). Radioactive isotope labeling studies suggest that most Rutaceae limonoids are derived from a nomilin-type intermediate (*seco*-A,D ring scaffolds), whereas Meliaceae limonoids are derived from an azadirone-type intermediate (intact A ring) (Fig. 1) (*4*, *5*, *18*, *19*). It is proposed that at least two main scaffold modifications are conserved in both plant families: a C-30 methyl shift of the protolimonoid scaffold (apo-rearrangement) and the conversion of the hemiacetal ring of melianol (**1**) to a mature furan ring with a concomitant loss of the C-25~C-28 carbon side chain (Fig. 1) (*20*). Additional modifications that are specific to Rutaceae and Meliaceae would then yield the nomilin- and azadirone-type intermediates. The diversity and array of protolimonoid structures that are isolated beyond melianol (**1**) (fig. S1) hint at a series of possible conserved biosynthetic transformations, including hydroxylation and/or acetoxylation on C-1, C-7, and C-21, which suggests the involvement of cytochrome P450s (CYPs), 2-oxoglutarate–dependent dioxygenases (2-ODDs), and acetyltransferases.

Despite extensive interest in the biology and chemistry of complex plant triterpenes over the past half century, few complete biosynthetic pathways have been described. A notable exception is the disease resistance saponin from oat, avenacin A-1, whose pathway consists of four CYP-mediated scaffold modifications and six side-chain tailoring steps (*21*). Barriers to pathway reconstitution of complex triterpenes include a lack of knowledge of the structures of key intermediates, the order of scaffold modification steps, the instability of pathway precursors, and the challenge of identifying candidate genes for the anticipated >10 enzymatic transformations that are required to generate advanced intermediates. Limonoids are no exception; to date, only the first three enzymatic steps to the protolimonoid melianol (**1**) from the primary metabolite 2,3-oxidosqualene have been elucidated (Fig. 1) (*20*). In this work, we used systematic transcriptome and genome mining and phylogenetic and homologous analysis, coupled with *Nicotiana benthamiana* as a heterologous expression platform, to identify suites of candidate genes from *Citrus sinensis* and *Melia azedarach* that can be used to reconstitute limonoid biosynthesis.

Identification of candidate limonoid biosynthetic genes

One genome of Rutaceae plants (*C. sinensis* var. Valencia) and several transcriptome resources, including from *Citrus* and Meliaceae plants (two from *A. indica* and one from *M. azedarach*), were previously used to identify the first three enzymes in the limonoid pathway (*20*). These included an oxidosqualene cyclase (*Cs*OSC1 from *C. sinensis*, *Ai*OSC1 from

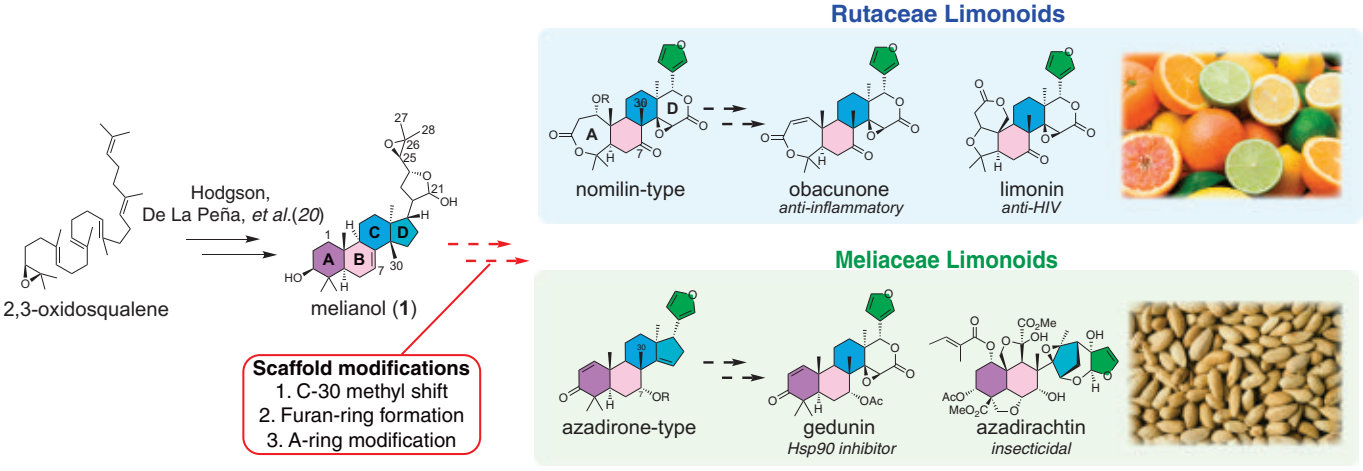


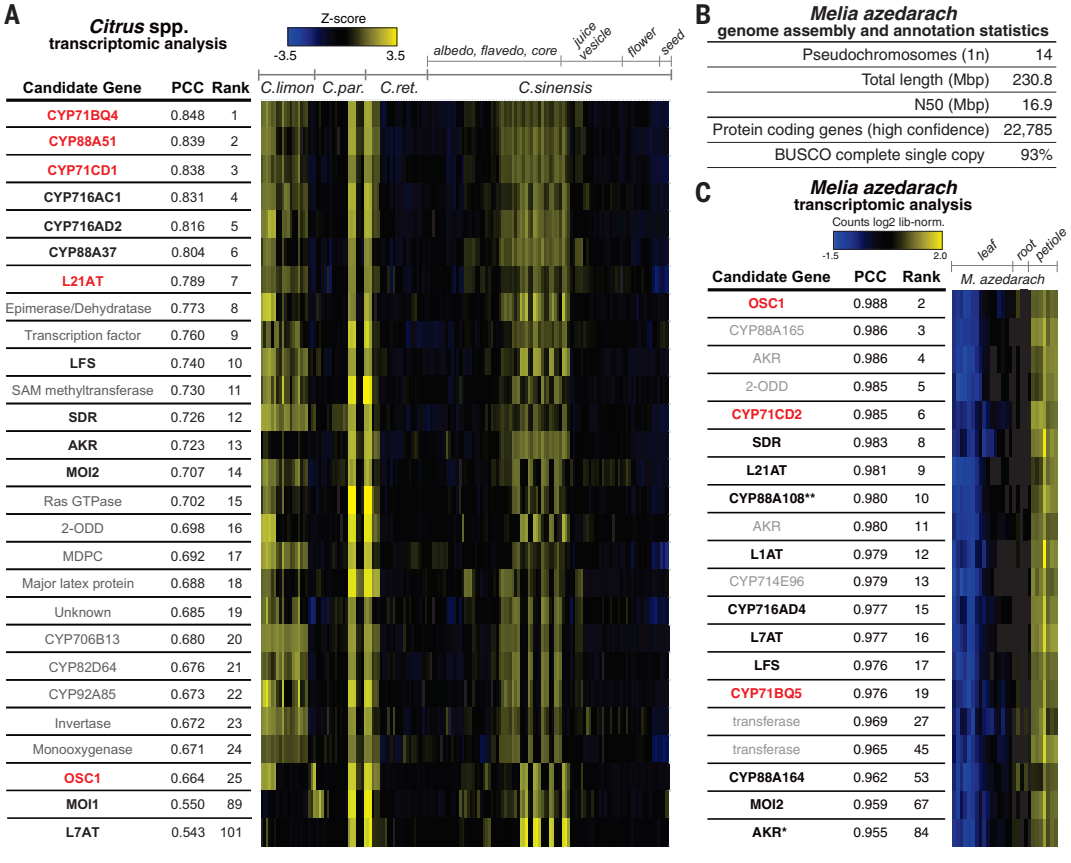
Fig. 1. Structures of Rutaceae and Meliaceae limonoids and proposed biosynthetic pathway. We previously characterized three conserved enzymes from both *Citrus* and *Melia* species that catalyze the formation of the protolimonoid melianol (**1**) from 2,3-oxidosqualene (*20*). Additionally, conserved scaffold modifications such as C-30 methyl shift, furan-ring formation, and A-ring modification are proposed to convert protolimonoids to true limonoids. Beyond this, Rutaceae limonoids differ from Meliaceae limonoids in two key structural features: *seco*-A,D ring and C-7 modification, which are proposed to be the result of Rutaceae- and Meliaceae-specific modifications. Exceptions to this rule could potentially arise from late-stage species-specific tailoring (fig. S43). Rutaceae limonoids are derived from nomilin-type intermediates, whereas Meliaceae limonoids are proposed to originate from azadirone-type intermediates. Although the exact point of pathway divergence is unknown, comparative analysis of the various protolimonoid structures suggests that C-1, C-7, and C-21 hydroxylation and/or acetoxylation are part of the conserved tailoring process. Obacunone and limonin are commonly found in various *Citrus* species and are responsible for the bitterness of their seeds. Azadirachtin (the most renowned Meliaceae limonoid) accumulates at high levels in the seeds of the neem tree, which are the source of commercial neem biopesticides. R represents H or Ac. Ac, acetyl; Me, methyl. [Photo credits: IgorDutina/iStock (top right); JIC photography (bottom right)]

A. indica, and *Ma*OSC1 from *M. azedarach*), and two CYPs (*Cs*CYP71CD1/*Ma*CYP71CD2 and *Cs*CYP71BQ4/*Ma*CYP71BQ5) that complete the pathway to melianol (*20*). To identify enzymes that further tailor melianol (**1**), we expanded our search to include additional sources. For Rutaceae enzyme identification, we included publicly available microarray data compiled by the Network Inference for Citrus Co-Expression (NICCE) (*22*). For Meliaceae enzyme identification, we generated additional RNA sequencing (RNA-seq) data and a reference-quality genome assembly and annotation.

Of publicly available microarray data for *Citrus*, fruit datasets were selected for in-depth analysis because *CsOSC1* expression levels were highest in the fruit and it has been implicated as the site of limonin biosynthesis and accumulation (*19*). Gene coexpression analysis was first performed on the *Citrus* fruit dataset by using only *CsOSC1* as the bait gene. This revealed promising candidate genes that exhibited highly correlated expression with *CsOSC1* (fig. S2). As we characterized more limonoid biosynthetic genes (as described below), we also included these as bait genes to enhance the stringency of coexpression analysis and further refine the candidate list. The top-ranking candidate list is rich in genes that are typically associated with secondary metabolism (Fig. 2A). The list specifically included multiple predicted CYPs, 2-ODDs, and acetyltransferases, which is consistent with the proposed biosynthetic transformations.

Efforts to identify and clone candidate genes from *M. azedarach* have previously been limited by the lack of a reference genome with high-quality gene annotations and by the lack of suitable transcriptomic data for coexpression analysis (i.e., multiple tissues, with replicates). Therefore, in parallel to our search in *Citrus*, we generated genomic and transcriptomic resources for *M. azedarach*. A pseudochromosome-level reference-quality *M. azedarach* genome assembly was generated by using PacBio long-read and Hi-C sequencing technologies (fig. S3 and table S1). Although the assembled genome size [230 million base pairs (Mbp)] is smaller than available literature predictions for this species of 421 Mbp (*23*), the chromosome number (1n = 14) matches literature reports (*23*) and was confirmed by karyotyping (fig. S4). The genome assembly annotation predicted 22,785 high-confidence protein coding genes (Fig. 2B and table S1). Benchmarking Universal Single-Copy Orthologs (BUSCO) assessment (*24*) of this annotation confirmed the completeness of the genome, because 93% of expected orthologs are present as complete single-copy genes (comparable with 98% in the gold standard *Arabidopsis thaliana*) (Fig. 2B and table S1).

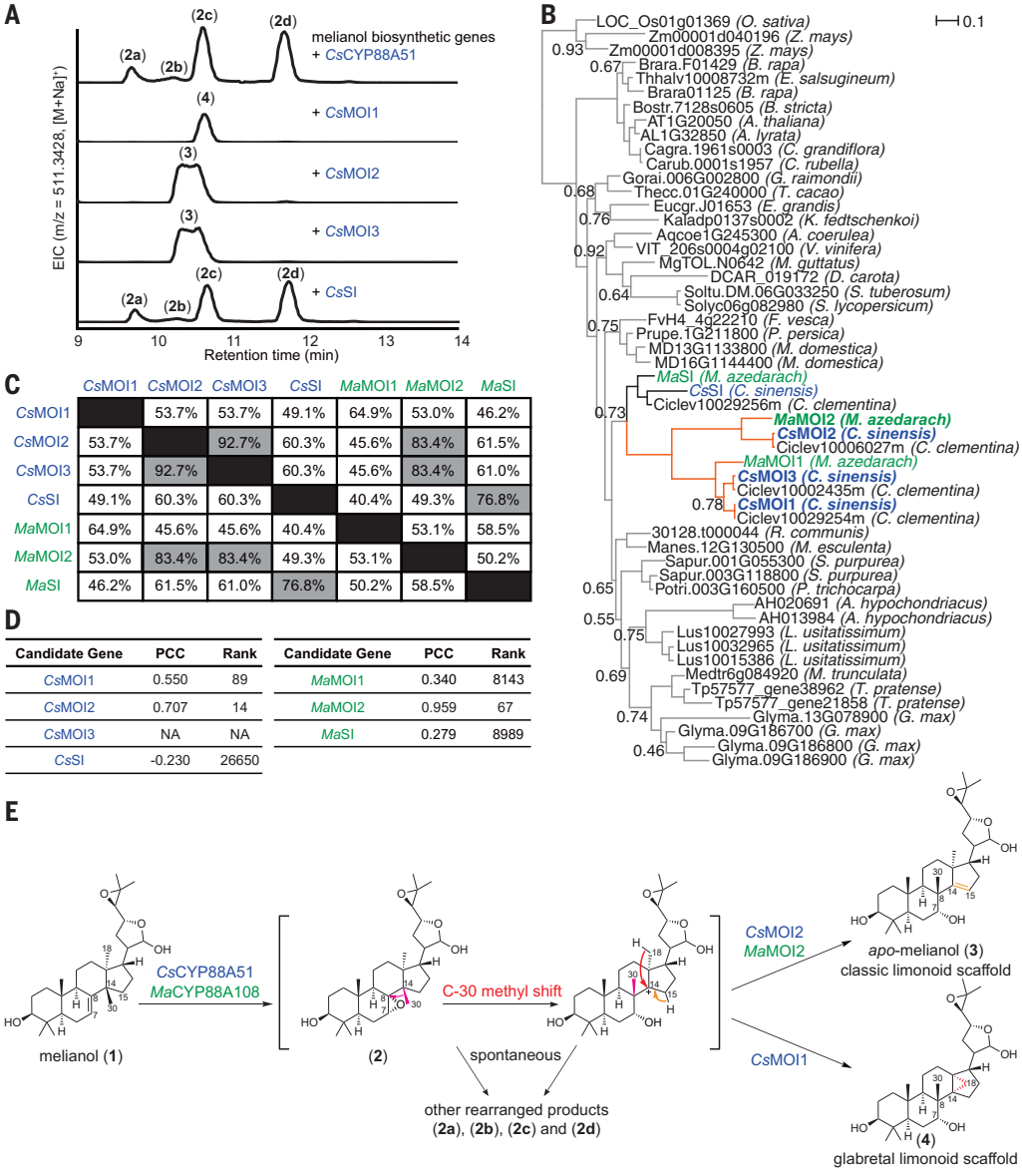
Illumina paired-end RNA-seq reads were generated for three different *M. azedarach* tissues (seven different tissues in total, with four replicates of each tissue) (table S2),



Candidate Gene	PCC	Rank
CYP71BQ4	0.848	1
CYP88A51	0.839	2
CYP71CD1	0.838	3
CYP716AC1	0.831	4
CYP716AD2	0.816	5
CYP88A37	0.804	6
L21AT	0.789	7
Epimerase/Dehydratase	0.773	8
Transcription factor	0.760	9
LFS	0.740	10
SAM methyltransferase	0.730	11
SDR	0.726	12
AKR	0.723	13
MOI2	0.707	14
Ras GTPase	0.702	15
2-ODD	0.698	16
MDPC	0.692	17
Major latex protein	0.688	18
Unknown	0.685	19
CYP706B13	0.680	20
CYP82D64	0.676	21
CYP92A85	0.673	22
Invertase	0.672	23
Monooxygenase	0.671	24
OSC1	0.664	25
MOI1	0.550	89
L7AT	0.543	101

Candidate Gene	PCC	Rank
OSC1	0.988	2
CYP88A165	0.986	3
AKR	0.986	4
2-ODD	0.985	5
CYP71CD2	0.985	6
SDR	0.983	8
L21AT	0.981	9
CYP88A108**	0.980	10
AKR	0.980	11
L1AT	0.979	12
CYP714E96	0.979	13
CYP716AD4	0.977	15
L7AT	0.977	16
LFS	0.976	17
CYP71BQ5	0.976	19
transferase	0.969	27
transferase	0.965	45
CYP88A164	0.962	53
MOI2	0.959	67
AKR*	0.955	84

Fig. 2. Genomic and transcriptomic analysis of *Citrus* and *Melia* resources. (**A**) Coexpression analysis of *C. sinensis* publicly available microarray expression data from NICCE (*22*) using *CsOSC1*, *CsCYP71CD1*, *CsCYP71BQ4*, *CsCYP88A51*, and *CsL21AT* as bait genes. Linear regression analysis was used to rank the top-25 genes on the basis of Pearson's correlation coefficient (PCC) to the bait genes of interest. The heat map displays the *Z*-score calculated from $\log_2$-normalized expression across the fruit dataset. The reported PCC value corresponds to the average value calculated by using each bait gene. Genes in red indicate bait genes used in analysis, and genes in black are functional limonoid biosynthetic genes (table S18). Functional candidates outside of the top-25 genes are also included. For identification of individual bait genes used in this analysis, see fig. S2. L1AT, limonoid C-1-*O*-acetyltransferase; L21AT, limonoid C-21-*O*-acetyltransferase; L7AT, limonoid C-7-*O*-acetyltransferase; OSC, oxidosqualene cyclase. (**B**) Summary of *M. azedarach* pseudochromosome genome assembly and annotation statistics (figs. S3 and S4 and tables S1 and S2). (**C**) Expression pattern of *M. azedarach* limonoid candidate genes selected on the basis of PCC to melianol biosynthetic genes [*MaOSC1*, *MaCYP71CD2*, and *MaCYP71BQ5* (*20*), shown in red] and biosynthetic annotation. The heat map [constructed using Heatmap3 V1.1.1 (*44*), with scaling by row (gene)] includes genes that are ranked within the top 87 for coexpression and are annotated with one of six InterPro domains of biosynthetic interest [IPR005123 (oxoglutarate/iron-dependent dioxygenase), IPR020471 (AKR), IPR002347 (SDR), IPR001128 (CYP), IPR003480 (transferase), and IPR007905 (emopamil-binding protein)]. Genes shown in black are newly identified functional limonoid biosynthetic genes (this study) (table S10). *Full-length gene identified in transcriptomic rather than genomic data through sequence similarity to *Cs*AKR (tables S10 and S19); **Gene previously identified as homolog of limonoid coexpressed gene from *A. indica* (*20*).



	CsMOI1	CsMOI2	CsMOI3	CsSI	MaMOI1	MaMOI2	MaSI
CsMOI1		53.7%	53.7%	49.1%	64.9%	53.0%	46.2%
CsMOI2	53.7%		92.7%	60.3%	45.6%	83.4%	61.5%
CsMOI3	53.7%	92.7%		60.3%	45.6%	83.4%	61.0%
CsSI	49.1%	60.3%	60.3%		40.4%	49.3%	76.8%
MaMOI1	64.9%	45.6%	45.6%	40.4%		53.1%	58.5%
MaMOI2	53.0%	83.4%	83.4%	49.3%	53.1%		50.2%
MaSI	46.2%	61.5%	61.0%	76.8%	50.2%	58.5%	

Candidate Gene	PCC	Rank
CsMOI1	0.550	89
CsMOI2	0.707	14
CsMOI3	NA	NA
CsSI	-0.230	26650

Candidate Gene	PCC	Rank
MaMOI1	0.340	8143
MaMOI2	0.959	67
MaSI	0.279	8989

Fig. 3. Characterization of MOIs. (**A**) Characterization of products generated through overexpression of MOIs and SIs by using transient gene expression in *N. benthamiana*. LC-MS extracted ion chromatograms (EICs) that resulted from overexpression of *At*HMGR, *Cs*OSC1, *Cs*CYP71CD1, *Cs*CYP71BQ4, *Cs*CYP88A51, and *Cs*MOIs and *Cs*SI in *N. benthamiana*. Representative EICs are shown (n = 3). (**B**) Phylogenetic tree (Bayesian) of SI genes from high-quality plant genomes. SI sequences from 33 plant species were identified and downloaded from Phytozome through pFAM assignments (PF05241). Branch supports are provided (excluding those >0.95), and monocot SIs have been used as an outgroup. Enzymes that have melianol oxide isomerase activity when tested by *Agrobacterium*-mediated expression in *N. benthamiana* with melianol (**1**) biosynthetic genes and *Cs*CYP88A51 or *Ma*CYP88A108 have been renamed MOI, for example, *Cs*MOI1-3 and *Ma*MOI2. Characterized MOIs from *C. sinensis* and *M. azedarach* that were selected for further analysis are bolded, and their respective tree branches are indicated in orange. Genes from *Citrus* are shown in blue, and those from *Melia* are shown in green. All species listed are identified in the materials and methods section of the supplementary materials. (**C**) Percentage protein identity of MOIs and SIs from *C. sinensis* and *M. azedarach*; those with sequence similarity >75% are highlighted in gray. (**D**) Coexpression of MOIs and SIs from *C. sinensis* and *M. azedarach* displaying rank and PCC as outlined in Fig. 2, A and C. (**E**) Proposed mechanism of *Cs*CYP88A51/*Ma*CYP88A108, *Cs*MOI2/*Ma*MOI2, and *Cs*MOI1. *Cs*CYP88A51/*Ma*CYP88A108 first oxidizes the C-7, C-8 position of melianol (**1**) to yield an unstable epoxide intermediate **2**, which can undergo spontaneous C-30 methyl shift from C-14 to C-8 (highlighted in red). Either **2** or the methyl-shifted product spontaneously forms a series of oxidized products (**2a** to **2d**). In the presence of MOIs, the rearrangement of **2** is guided to form either **3** or **4**, and **2a** to **2d** are not observed. The structures of **2a** to **2d** are not determined, but their MS fragmentation patterns suggest that they are isomeric molecules that result from a single oxidation of melianol (**1**), which does not exclude the possibility them of being **2**, **3**, or **4** [as shown for *Ailanthus altissima* CYP71BQ17 (*35*)].

which were previously shown to differentially accumulate and express limonoids and their biosynthetic genes (*20*). Read counts were generated by aligning RNA-seq reads to the genome annotation, and EdgeR (*25*) was used to identify a subset of 18,151 differentially expressed genes ($P < 0.05$). The known melianol biosynthetic genes *MaOSC1*, *MaCYP71CD2*, and *MaCYP71BQ5* (*20*) were used as bait genes for coexpression analysis across the sequenced tissues, and the resulting ranked list was filtered by their InterPro domain annotations to enrich for relevant biosynthetic enzyme-coding genes. This informed the selection of 17 candidate genes for further investigation for functional analysis along with *Citrus* candidates (Fig. 2C).

Citrus CYP88A51 and *Melia* CYP88A108 act with different melianol oxide isomerases to form distinct protolimonoid scaffolds

Top-ranking genes from both the *Citrus* and *Melia* candidate lists (Fig. 2, A and C) were tested for function by *Agrobacterium*-mediated transient expression in *N. benthamiana* with the previously reported melianol (**1**) biosynthetic enzymes *Cs*OSC1, *Cs*CYP71CD1, and *Cs*CYP71BQ4 or *Ai*OSC1, *Ma*CYP71CD2, and *Ma*CYP71BQ4. Liquid chromatography–mass spectrometry (LC-MS) analysis of crude methanolic extracts from *N. benthamiana* leaves revealed that the expression of either *Cs*CYP88A51 or *Ma*CYP88A108, in combination with their respective melianol biosynthesis genes, led to the disappearance of melianol (**1**) and the accumulation of multiple mono-oxidized

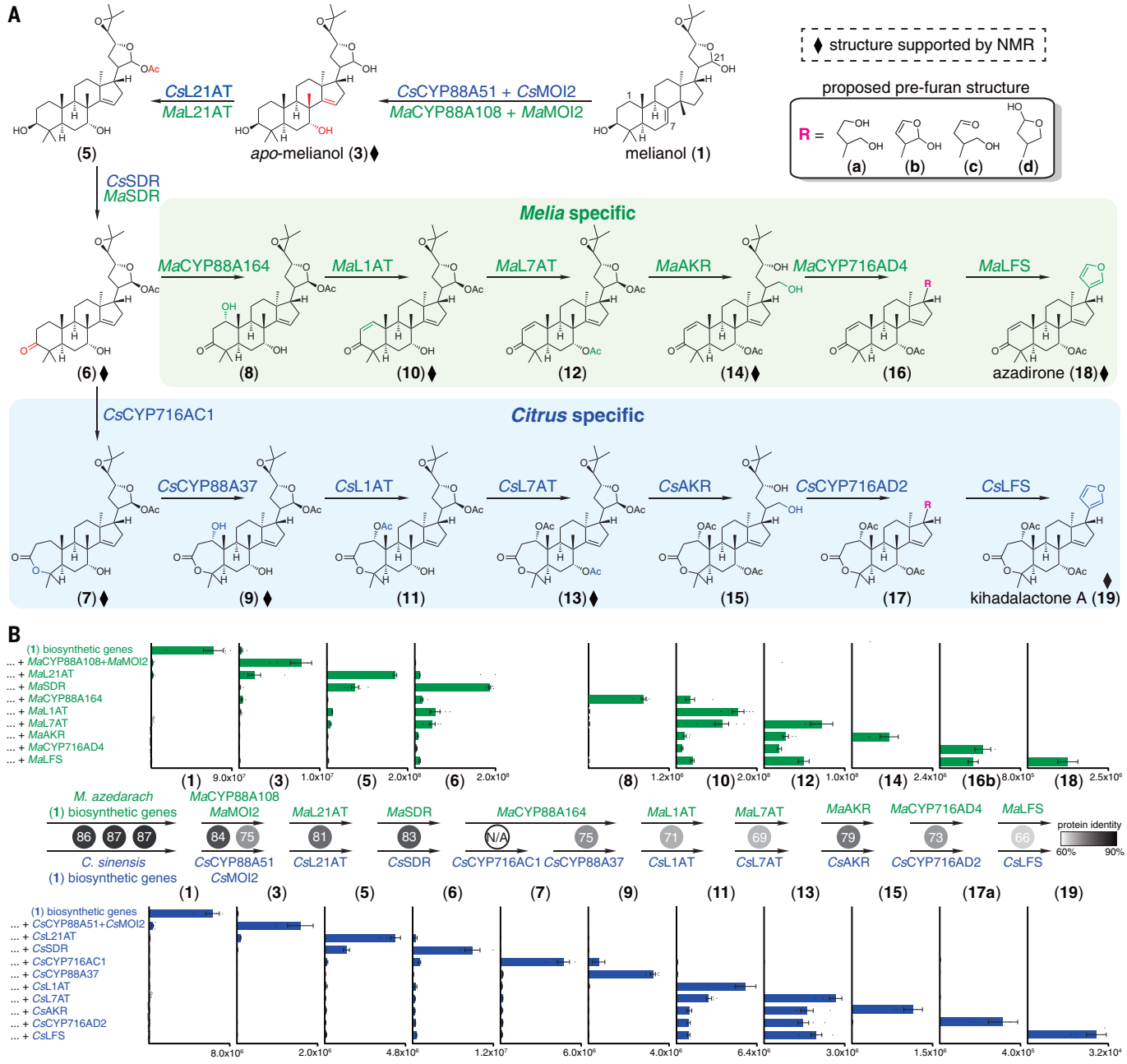


Fig. 4. Complete biosynthetic pathway to azadirone (18) and kihadalactone A (19). (**A**) Gene sets that lead to the production of azadirone (**18**) and kihadalactone A (**19**) in *N. benthamiana* leaves. Genes from *Citrus* are shown in blue, and those from *Melia* are shown in green. The arrow reflects accumulation of the metabolites after addition of the associated enzyme as shown in (B) rather than true enzymatic substrate-product relationship. In addition, limonoid biosynthesis likely proceeds as a network; other possible reaction sequences are shown in fig. S40. Diamonds represent intermediates whose structures were supported either by NMR analysis of the purified product or comparison with an authentic standard (**18**). Compounds **3**, **6**, **9**, **10**, **13**, and **14** were purified from *N. benthamiana* leaf extracts that expressed the respective biosynthetic gene sets and were analyzed by NMR; the structures of **7** and **19** are supported by partial NMR. Additionally, a side product, **20**, that formed in experiments with all pathway enzymes up to and including *Ma*CYP716AD4 but without *Ma*L7AT (fig. S44) was purified and confirmed by NMR (table S20); similar activity was observed for *Cs*CYP716AD2 (fig. S45 and supplementary text). (**B**) Integrated peak area of EIC for each pathway that intermediates produced in *N. benthamiana* after sequential coexpression of individual enzymes. Values and error bars represent the mean and the standard error of the mean; $n = 6$ biological replicates. Numbered circles with grayscale coloration indicate the percentage identity between homologous proteins. Compound **1** biosynthetic genes comprise *Ma*OSC1/*Cs*OSC1, *Ma*CYP71CD2/*Cs*CYP71CD1, and *Ma*CYP71BQ5/*Cs*CYP71BQ4. *Cs*CYP88A37 is a homolog of *Ma*CYP88A164, whereas *Cs*CYP716AC1 has no *Melia* homolog.

products (Fig. 3A and figs. S5 and S6). This result suggested that although these CYP88A enzymes accept melianol as a substrate, the resulting products could be unstable or undergo further modification by endogenous *N. benthamiana* enzymes.

Despite the accumulation of multiple related metabolites, we continued to screen additional coexpressed candidate genes for further activity. This screen included homologs of *A. thalianaHYDRA1*, which encodes an endoplasmic reticulum membrane protein known

as a sterol isomerase (SI) (two from the *Citrus* candidate list and one from the *Melia* list). SIs are exclusively associated with phytosterol and cholesterol biosynthesis, in which they catalyze double-bond isomerization from the C-8 to the C-7 position. They are present in all domains of life and are required for normal development of mammals (*26*), plants (*27*), and yeast (*28*). Testing of these putative SIs through transient *Agrobacterium*-mediated gene expression in *N. benthamiana* resulted in a marked change of the metabolite profile with the accumulation of a single mono-oxidized product with no mass change (Fig. 3A and fig. S7). We suspected that these enzymes were able to capture unstable intermediates and promote isomerization of the C-30 methyl group that is required to generate mature limonoids. These SIs are therefore renamed melianol oxide isomerases (MOIs), *CsMOI1* to *CsMOI3* and *MaMOI2*, because of their ability to generate isomers of mono-oxidized melianol products.

SIs are typically found as single-copy genes in given plant species. Surprisingly, we found additional putative SI genes in the *C. sinensis* and *M. azedarach* genomes: four and three, respectively (fig. S8). Phylogenetic analysis of SIs across a set of diverse plant species revealed that SIs from *C. sinensis* and *M. azedarach* fall into two distinct subclades (Fig. 3B). The more conserved of these clades contained one sequence from each species (*CsSI* and *MaSI*), whereas the more divergent clade contained the remaining SIs (*CsMOI1* to *CsMOI3* and *MaMOI1* and *MaMOI2*). This suggested that *CsSI* and *MaSI* are the conserved genes involved in phytosterol biosynthesis. Comparison of all *C. sinensis* and *M. azedarach* SI and MOI protein sequences showed that *Cs*MOI2 is ~93% identical at the protein level to *Cs*MOI3 and ~83% to *Ma*MOI2, but only ~54 and ~60% similar to *Cs*MOI1 and *Cs*SI, respectively (Fig. 3C). Although *CsMOI1*, *CsMOI2*, and *MaMOI2* ranked among the top 100 genes in our coexpression analysis lists (Fig. 3D), *CsSI*, *MaMOI1*, and *MaSI* do not coexpress with limonoid biosynthetic genes. The absence of *CsMOI3* from this list is attributed to the lack of specific microarray probes required for expression monitoring. Notably, screening of *Cs*SI in the *N. benthamiana* expression system did not change the product profile of *Cs*CYP88A51, which is consistent with its predicted involvement in primary metabolism on the basis of the phylogenetic analysis (Fig. 3A).

To determine the chemical structures of the isomeric products formed through the action of these MOIs, we carried out large-scale expression experiments in *N. benthamiana* and isolated 13.1 mg of pure product. Nuclear magnetic resonance (NMR) analysis revealed the product of *Ma*MOI2 to be the epimeric mixture apo-melianol (**3**), which bears the characteristic limonoid scaffold with a migrated C-30 methyl group on C-8, a C-14/15 double bond, and C-7 hydroxylation (Fig. 3E and table S3) (*29*). Although the structure of the direct product of *Cs*MOI2 was not determined until after the discovery of two additional downstream tailoring enzymes, NMR analysis also confirmed C-8 methyl migration (table S4). These data indicate that, as predicted by sequence analysis, *Cs*MOI2 and *Ma*MOI2 indeed are functional homologs and catalyze a key step in limonoid biosynthesis by promoting a previously uncharacterized methyl shift. Analysis of the product that formed with the expression of *Cs*MOI1 indicated the presence of a metabolite with a different retention time relative to apo-melianol (**3**) (Fig. 3A). Isolation and NMR analysis of **4′**, a metabolite that is derived from **4** after the inclusion of two additional tailoring enzymes (table S5), indicated C-30 methyl group migration to C-8 and cyclopropane ring formation through bridging of the C-18 methyl group to C-14.

On the basis of the characterized structures, we propose that in the absence of MOIs, the CYP88A homologs form the unstable C-7/8 epoxide **2**, which may either spontaneously undergo a Wagner-Meerwein rearrangement via C-30 methyl group migration and subsequent epoxide-ring-opening or degrade through other routes to yield multiple rearranged products **2a** to **2c** and **3** (Fig. 3E). MOIs appear to stabilize the unstable carbocation intermediate and isomerize it to two types of limonoids: *Cs*MOI2, *Cs*MOI3, and *Ma*MOI2 form the C-14/15 double bond scaffold (classic limonoids), whereas *Cs*MOI1 forms the cyclopropane ring scaffold (glabretal limonoids). Glabretal limonoids have been isolated from certain Meliaceae and Rutaceae species before but are less common (*30*, *31*). *Cs*CYP88A51, *Ma*CYP88A108, and two different types of MOIs are thus responsible for rearrangement from melianol (**1**) to either **3** or **4** through the epoxide intermediate **2**. These MOIs represent the neofunctionalization of SIs from primary metabolism in plants.

Characterization of conserved tailoring enzymes L21AT and SDR

Having identified enzymes for the methyl shift present in the limonoids, we continued screening other candidate genes (Fig. 2, A and C) for activity on **3** toward downstream products. BAHD-type acetyltransferases (named *Cs*L21AT or *Ma*L21AT, limonoid 21-*O*-acetyltransferse) and short-chain dehydrogenase reductases (SDRs) (*Cs*SDR and its homolog *Ma*SDR) result in the loss of compound **3** and the accumulation of acetylated products and a dehydrogenated product, respectively (figs. S9 to S12). Although the sequence of events can be important for some enzymatic transformations in plant biosynthesis, L21AT and SDR homologs appear to have broad substrate specificity. Our data suggest that L21AT can act on **1** or **3** and that SDR is active on all intermediates after the OSC1 product (figs. S13 and S14), which suggests a flexible reaction order in the early biosynthetic pathway.

Furthermore, the products formed from the modification of **3** by both *Citrus* and *Melia* L21AT and SDR homologs were purified by large-scale *N. benthamiana* expression and structurally determined by NMR to be 21-(*S*)-acetoxyl-apo-melianone (**6**) (Fig. 4A, fig. S15, and tables S4 to S7). Compound **6** is a protolimonoid previously purified from the Meliaceae species *Chisocheton paniculatus* (*32*) and is also detectable in *M. azedarach* tissues (fig. S16). L21AT likely stereoselectively acetylates the 21-(*S*) isomer; a possible role for this transformation is stabilization of the hemiacetal ring that is observed as an epimeric mixture in melianol (**1**) (*20*) and apo-melianol (**3**) (table S3). Overall, our results indicate that L21AT acetylates the C-21 hydroxyl and SDR oxidizes the C-3 hydroxyl to the ketone on early protolimonoid scaffolds.

Citrus and *Melia* CYPs catalyze distinct limonoid A-ring modifications

Further *Citrus* and *Melia* candidate screens (Fig. 2, A and C) support the activity of two *Citrus* CYPs, *Cs*CYP716AC1 and *Cs*CYP88A37, that are each capable of oxidizing **6** directly to **7** and **8** or consecutively to **9** (Fig. 4A and figs. S17 to S19) and that one CYP from *Melia* (*Ma*CYP88A164, a homolog of *Cs*CYP88A37) is also capable of oxidizing **6** to **8** (Fig. 4A and fig. S20). Purification and NMR analysis of the downstream product **9** revealed it to be 1-hydroxy-luvungin A, which bears an A-ring lactone (table S8). Additional NMR product characterization suggests that *Cs*CYP716AC1 is responsible for A-ring lactone formation and *Cs*CYP88A37 is responsible for C-1 hydroxylation (table S9). Although the exact order of oxidation steps to **9** appeared to be interchangeable for *Cs*CYP716AC1 and *Cs*CYP88A37, the incomplete disappearance of **6** by *Cs*CYP88A37 suggests that oxidation by *Cs*CYP716AC1 takes precedence (fig. S19).

In the absence of *Cs*SDR, neither *Cs*CYP716AC1 nor *Cs*CYP88A37 result in an oxidized protolimonoid scaffold, which suggests the necessary involvement of the C-3 ketone for further processing (fig. S21). These results, in combination with NMR characterization, indicated that *Cs*CYP716AC1 is likely responsible for Baeyer-Villiger oxidation to the A-ring lactone structure signature of Rutaceae limonoids. Comparative transcriptomics in *M. azedarach* revealed the lack of an obvious *Cs*CYP716AC1 homolog. The closest *Melia* enzyme to *Cs*CYP716AC1 is truncated, not coexpressed with melianol biosynthetic genes, and only shares 63% protein identity (table S10). These results highlight a branch point

between biosynthetic routes in the Rutaceae and Meliaceae families.

Acetylations complete tailoring in both *Citrus* and *Melia* protolimonoid scaffolds and set the stage for furan ring biosynthesis

Subsequent *Citrus* and *Melia* gene candidate screens (Fig. 2, A and C) revealed further activity of BAHD acetyltransferases. *Cs*L1AT and its homolog *Ma*L1AT (named limonoid 1-*O*-acetyltransferase) appear to be active on **9** and **8**, respectively (figs. S22 and S23). When *Cs*L1AT was coexpressed with the biosynthetic genes for **9**, a molecule **11** with mass corresponding to acetylation of **9** was observed. When *Cs*CYP88A37 was omitted, acetylation of **7** was not observed (fig. S24), which suggests that *Cs*L1AT acetylates the C-1 hydroxyl of **9** to yield **11**. However, when *Cs*CYP716AC1 was omitted from the *Citrus* candidates or when *Ma*L1AT was tested, the dehydration scaffold **10** accumulated (figs. S23 and S24). Large-scale transient plant expression, purification, and NMR analysis of the dehydration product showed that the structure **10** (tables S11 and S12) contains a C-1/2 double bond and is an epimer of a previously reported molecule from *A. indica* (*33*). Compound **10** also accumulates in *M. azedarach* extracts (fig. S16). Two more coexpressed *Citrus* and *Melia* acetyltransferase homologs, *Cs*L7AT and *Ma*L7AT (named limonoid 7-*O*-acetyltransferase), were found to result in acetylated scaffolds **12** and **13**; modification at the C-7 hydroxyl was confirmed by the purification and NMR analysis of **13** and its degradation product **13′** (Fig. 2, A and C; figs. S25 and S26; and tables S13 and S14).

Taken together, these data suggest that three acetyltransferases (L1AT, L7AT, and L21AT) act in the biosynthesis of the triacetylated 1,7,21-*O*-acetyl protolimonid (**13**) (Fig. 4A). However, we also observed the accumulation of two diacetylated intermediates, **11** (1,21-*O*-acetyl) and **11a** (1,7-*O*-acetyl), when testing gene sets that lead to accumulation of **13** (fig. S27). This observation hints at the possibility of multiple sequences for enzymatic steps that comprise a metabolic network, at least in the context of pathway reconstitution in the heterologous host *N. benthamiana*.

Downstream enzymes complete the biosynthesis to the furan-containing products azadirone (18) and kihadalactone A (19)

With acetylation established, the key enzymes that are involved in the C-4 scission that is implicated in furan ring formation still remained elusive. It was unclear which enzyme classes could catalyze these modifications. We screened gene candidates by means of combinatorial transient expression in *N. benthamiana* as previously described and ultimately identified three active candidate pairs (one from each species): the aldo-keto reductases (AKRs) (*Cs*AKR/*Ma*AKR), the CYP716ADs (*Cs*CYP716AD2/*Ma*CYP716AD4), and the 2-ODDs [named limonoid furan synthase (LFS), *Cs*LFS/*Ma*LFS] (Fig. 2, A and C). Systematic testing of these gene sets resulted in the accumulation of the furan-containing molecules azadirone (**18**) and kihadalactone A (**19**), two limonoids that are present in the respective native species. When *Cs*AKR/*Ma*AKR was tested alone in our screens, we identified the appearance of a peak with mass corresponding to reductive deacetylation of **12** or **13** (figs. S28 and S29). The product generated by expression of the *Melia* gene set in *N. benthamiana* was purified and characterized by NMR analysis to be the 21,23-diol **14** (Fig. 4A and table S15). Thus, the corresponding *Cs*AKR product **15** was proposed to share the same diol motif.

Transient expression of *Ma*CYP716AD4 or *Cs*CYP716AD2 with the biosynthetic genes for **14** or **15** resulted in two pairs of peaks, each with C-4 loss. Proposed structures indicate a C_4H_6O fragment loss (**16a** and **17a**) and a $C_4H_{10}O$ fragment loss (**16b** and **17b**) from their respective precursors (Fig. 4A and figs. S30 and S31). It is unclear whether these observed masses correspond to the true products of CYP716ADs or whether these are further modified by endogenous *N. benthamiana* enzymes. CYP716AD products are proposed to contain C-21 hydroxyl and C-23 aldehyde functionalities (**16c** and **17c**), which could also spontaneously form the five-membered hemiacetal ring (**16d** and **17d**) (Fig. 4A and fig. S32). A peak with a mass equivalent to **16c** or **16d** is identifiable alongside **16a** and **16b** when transiently expressing *Ma*CYP716AD4 with the biosynthetic genes required for accumulation of **14** (fig. S31). We found that additional coexpression of LFS with the characterized genes that result in **16** and **17** yields accumulation of products **18** and **19** (figs. S33 and S34). On the basis of the predicted chemical formula, MS fragmentation pattern, and NMR analysis (fig. S33 and table S16), we proposed the product of *Cs*LFS to be kihadalactone A (**19**), a known furan-containing limonoid (*34*) that was previously identified in extracts from the Rutaceae plant *Phellodendron amurense*. We detected the presence of **19** in *P. amurense* seed samples (fig. S35), which confirmed the prior reports of accumulation. Similarly, when *Ma*LFS was included in the coexpression, a product with a mass equivalent to the furan-containing limonoid azadirone (**18**) was observed (fig. S34). The production of azadirone (**18**) in *N. benthamiana* was confirmed by comparison with an analytical standard (isolated from *A. indica* leaf powder and analyzed by NMR) (fig. S36 and table S17). In addition, we detected azadirone in extracts from three Meliaceae species (fig. S36).

Taken together, we have discovered the 10- and 11-step biosynthetic transformations that enable a reconstitution of the biosynthesis of two known limonoids, azadirone (**18**) and kihadalactone A (**19**), as well as an enzyme that catalyzes the formation of the alternative glabretal scaffold (*Cs*MOI1). Sequential introduction of these enzymes into *N. benthamiana* transient coexpression experiments demonstrate stepwise transformations that lead to **18** and **19** (Fig. 4B). All of the enzymes involved in the biosynthesis of **18** and **19**, except *Cs*CYP716AC1, are homologous pairs and show a gradual decreasing trend in protein identity from 86% for the first enzyme pair *Cs*OSC1/*Ma*OSC1 to 66% for *Cs*LFS/*Ma*LFS. Despite the varied protein identities (Fig. 4B), these homologous enzymes from *Melia* or *Citrus* can be used to create functional hybrid pathways that comprise a mix of species genes, which supports a promiscuous evolutionary ancestor for each of the limonoid biosynthetic enzymes (fig. S37).

Discussion

A major challenge in elucidating pathways that involve many (e.g., >10) enzymatic steps is to determine whether the observed enzymatic transformations in a heterologous host are "on pathway" and, if so, in what order they occur. It is important to note that although all enzymes described in Fig. 4 play a role in the production of final limonoid products, the sequence of enzymatic steps shown by the arrows is proposed on the basis of the accumulation of observed metabolites after the addition of each enzyme in the *N. benthamiana* heterologous expression system, and other sequences of steps are possible. For example, we have shown that *Cs*AKR likely does not accept hemiacetal (**13**) directly as a substrate (fig. S38), despite our observation that it accumulates as a major metabolite when all upstream enzymes are expressed. Although one expects a pathway without *Cs*L21AT to still be functional because the C-21 acetal product (**11a**) appears to undergo reduction by *Cs*AKR to yield **15**, attempts to drop out *Cs*L21AT led to a substantially reduced yield of **19** (fig. S39), which suggests that *Cs*L21AT might have other unexpected roles in the pathway. In addition, reconstitution of several partial pathways indicates that some pathway enzymes can accept multiple related substrates. For example, each step after apo-melianol can diverge into multiple pathways, which is likely due to the promiscuity of these enzymes. Taken together, these data indicate that enzymes in limonoid biosynthesis might collectively function as a metabolic network (fig. S40). Further study of each individual enzyme in vitro with purified substrate will be required to quantify substrate preference. This metabolic network that we observed in *N. benthamiana* suggests

one possible strategy for how Rutaceae species access such a diverse range of limonoids; we anticipate that additional enzymes will further expand the network, for example, for the oxidative cleavage of ring C, which ultimately results in the most extensively rearranged and modified limonoid scaffolds isolated to date, for example, azadirachtin (Fig. 1).

Among the 12 chemical transformations that are catalyzed by the 22 enzymes that we characterized in this study, several are not previously known in plant-specialized metabolism. For example, MOI1 and MOI2, which appear to have evolved from SIs, are capable of catalyzing two different scaffold rearrangements despite their conserved active-site residues (fig. S41). The colocalization of the limonoid biosynthetic gene *MaMOI2* with two other nonlimonoid SI genes in the *M. azedarach* genome is consistent with the origin of *MaMOI2* by tandem duplication and neofunctionalization (fig. S42); this genomic arrangement is conserved in *Citrus* on chromosome 5 as well. Furthermore, recent findings demonstrate a similar role of these enzymes in quassinoid biosynthesis (*35*). Other noteworthy enzymatic reactions in the limonoid pathway include C-4 scission and furan ring installation, which generate an important pharmacophore of the limonoids. Although furan-forming enzymes have been reported from other plants (*36*–*38*), the AKR, CYP716AD, and 2-ODD module described here represents an alternative mechanism of furan formation via oxidative cleavage of a C-4 moiety. Along with the SIs (MOIs), the AKR and 2-ODDs add to the growing pool of enzyme families (*39*, *40*) associated with primary sterol metabolism that appear to have been recruited to plant secondary triterpene biosynthesis, which is likely due to the structural similarities between sterols and tetracyclic triterpenes.

Limonoids are only one of many families of triterpenes from plants with complex scaffold modifications. Other examples include the *Schisandra* nortriterpenes (*41*), quinonoids (*42*), quassinoids (*43*), and dichapetalins (*42*); each represent a large collection of structurally diverse terpenes that contain several members with potent demonstrated biological activity but no biosynthetic route. Despite the value of these complex plant triterpenes, individual molecular species are typically only available through multistep chemical synthesis routes or isolation from producing plants, which limits drug development (*15*) and agricultural utility (*9*). Many are only easily accessible in an unpurified extract form that contains multiple chemical constituents; for example, azadirachtin, one of the most potent limonoids, can only be obtained commercially as a component of neem oil. Our results demonstrate that pathways to triterpenes with complex scaffold modifications can be reconstituted in a plant host, and the gene sets we describe enable rapid production and isolation of naturally occurring limonoids. We anticipate that bioproduction of limonoids will serve as an attractive method to generate clinical candidates for evaluation and that stable engineering of the limonoid pathway could be a viable strategy for sustainable crop protection.

REFERENCES AND NOTES

1. Y. Y. Zhang, H. Xu, *RSC Advances* **7**, 35191–35220 (2017).
2. Q.-G. Tan, X.-D. Luo, *Chem. Rev.* **111**, 7437–7522 (2011).
3. A. Roy, S. Saraf, *Biol. Pharm. Bull.* **29**, 191–201 (2006).
4. F. Mulani, S. Nandikol, H. Thulasiram, ChemRxiv 2022-2bpb9 [Preprint] (2022); https://doi.org/10.26434/chemrxiv-2022-2bpb9.
5. J. Luo, Y. Sun, Q. Li, L. Kong, *Nat. Prod. Rep.* **39**, 1325–1365 (2022).
6. S. Yamashita *et al*., *Angew. Chem. Int. Ed.* **54**, 8538–8541 (2015).
7. G. E. Veitch *et al*., *Angew. Chem. Int. Ed.* **46**, 7629–7632 (2007).
8. J. Li, F. Chen, H. Renata, ChemRxiv 2022-tqvbw [Preprint] (2022); https://doi.org/10.26434/chemrxiv-2022-tqvbw.
9. E. D. Morgan, *Bioorg. Med. Chem.* **17**, 4096–4105 (2009).
10. M. Puri, S. S. Marwaha, R. M. Kothari, J. F. Kennedy, *Crit. Rev. Biotechnol.* **16**, 145–155 (1996).
11. L. Battinelli *et al*., *Planta Med.* **69**, 910–913 (2003).
12. X. Luo *et al*., *Pharm. Biol.* **58**, 886–897 (2020).
13. J. Lamb *et al*., *Science* **313**, 1929–1935 (2006).
14. J. N. Spradlin *et al*., *Nat. Chem. Biol.* **15**, 747–755 (2019).
15. B. Tong *et al*., *ACS Chem. Biol.* **15**, 1788–1794 (2020).
16. V. Domingo, J. F. Arteaga, J. F. Quílez del Moral, A. F. Barrero, *Nat. Prod. Rep.* **26**, 115–134 (2009).
17. W. J. Baas, *Phytochemistry* **24**, 1875–1889 (1985).
18. P. Ou, S. Hasegawa, Z. Herman, C. H. Fong, *Phytochemistry* **27**, 115–118 (1988).
19. S. Hasegawa, in *Citrus Limonoids: Functional Chemicals in Agriculture and Foods*, M. A. Berhow, S. Hasegawa, G. D. Manners, Eds., vol. 758 of ACS Symposium Series (American Chemical Society, 2000), pp. 9–30.
20. H. Hodgson *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 17096–17104 (2019).
21. Y. Li *et al*., *Nat. Commun.* **12**, 2563 (2021).
22. D. C. J. Wong, C. Sweetman, C. M. Ford, *BMC Plant Biol.* **14**, 186 (2014).
23. D. Ohri, A. Bhargava, A. Chatterjee, *Plant Biol.* **6**, 555–561 (2004).
24. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, *Bioinformatics* **31**, 3210–3212 (2015).
25. M. D. Robinson, D. J. McCarthy, G. K. Smyth, *Bioinformatics* **26**, 139–140 (2010).
26. N. Braverman *et al*., *Nat. Genet.* **22**, 291–294 (1999).
27. M. Souter *et al*., *Plant Cell* **14**, 1017–1031 (2002).
28. S. Silve *et al*., *J. Biol. Chem.* **271**, 22434–22440 (1996).
29. W. Herz, H. Grisebach, G. W. Kirby, Ch. Tamm, Eds., *Fortschritte der Chemie organischer Naturstoffe: Progress in the Chemistry of Organic Natural Products*, vol. 46 (Springer, 1984).
30. G. Ferguson *et al*., *J. Chem. Soc. Chem. Commun.* **1973**, 159–196 (1973).
31. A.-R. Choi *et al*., *Chem. Pharm. Bull. (Tokyo)* **63**, 1065–1069 (2015).
32. J. Connolly, C. Labbé, D. S. Rycroft, D. A. H. Taylor, *J. Chem. Soc., Perkin Trans. 1* **1**, 2959–2964 (1979).
33. G. Chianese *et al*., *J. Nat. Prod.* **73**, 1448–1452 (2010).
34. K. Kishi, K. Yoshikawa, S. Arihara, *Phytochemistry* **31**, 1335–1338 (1992).
37. C. M. Bertea, M. Schalk, F. Karp, M. Maffei, R. Croteau, *Arch. Biochem. Biophys.* **390**, 279–286 (2001).
38. J.-J. Song *et al*., *Plant Physiol.* **188**, 1496–1506 (2022).
39. G. Polturak *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2123299119 (2022).
40. X. Qi *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18848–18853 (2006)
41. Y.-M. Shi, W.-L. Xiao, J.-X. Pu, H.-D. Sun, *Nat. Prod. Rep.* **32**, 367–410 (2015).
42. R.-Y. Kuo, K. Qian, S. L. Morris-Natschke, K.-H. Lee, *Nat. Prod. Rep.* **26**, 1321–1344 (2009).
43. Z. Guo, S. Vangapandu, R. W. Sindelar, L. A. Walker, R. D. Sindelar, *Curr. Med. Chem.* **12**, 173–190 (2005).
44. S. Zhao, Y. Guo, Q. Sheng, Y. Shyr, *BMC Bioinformatics* **15**, P16 (2014).

ACKNOWLEDGMENTS

We acknowledge J. Staples for performing the extractions of Meliaceae material sourced from Kew Gardens and S. Lynch (head of the NMR facility, Stanford University, Department of Chemistry) for helpful discussion on NMR structural elucidation. **Funding:** Work undertaken in E.S.S.'s laboratory was supported by the National Institutes of Health (NIH) grants NIH U01 GM110699 and NIH R01 GM121527. The work undertaken in A.O.'s laboratory was partly funded through a Biotechnology and Biological Sciences Research Council Industrial Partnership (BBSRC) Award with Syngenta (BB/T015063/1), and A.O.'s lab is also supported by the John Innes Foundation and the BBSRC Institute Strategic Programme Grant "Molecules from Nature – Products and Pathways" (BBS/E/J/00PR9790). **Author contributions:** R.D.L.P., H.H., A.O., and E.S.S. conceived of the project and, with assistance from J.C.-T.L., designed the research. H.H. generated and analyzed *M. azedarach* genome and transcriptome data. R.D.L.P. analyzed *Citrus* gene expression data and selected candidate genes from *Citrus* and, with J.C.T.L., expressed and characterized biosynthetic genes and metabolic products. L.E.J. assisted with isolation of *Citrus* intermediates. H.H. analyzed the *Melia* sequence resources and selected, expressed, and characterized *Melia* biosynthetic genes and metabolic products. J.C.T.L. and M.J.S. performed NMR analysis on the *Citrus* and *Melia* products, respectively. J.L.-M. advised on *M. azedarach* genomics. A.H. performed chromatin cross-linking and DNA extraction on *M. azedarach* tissues for Hi-C analysis by phase genomics. A.C.M. performed karyotyping on *M. azedarach* roots. C.O. combined the pseudochromosome-level genome assembly with the *M. azedarach* annotation and constructed the phylogenetic tree. R.D.L.P., J.C.T.L., H.H., A.O., and E.S.S. analyzed the data and wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All Citrus genes in this study have been deposited in the National Center for Biotechnology Information (NCBI) database under accession numbers OQ091238 to OQ091249. The *M. azedarach* genome has been deposited in the NCBI database (accession number PRJNA906622) along with the accompanying RNA-seq data (accession number PRJNA906055). Coding sequences for the functional *M. azedarach* genes described in this study have also been deposited in GenBank under accession numbers OP947595 to OP947604. **License information:** Copyright © 2023 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.adf1017
Materials and Methods
Supplementary Text
Figs. S1 to S45
Tables S1 to S24
References (*45*–*81*)
MDAR Reproducibility Checklist
Data S1

Submitted 29 September 2022; accepted 23 December 2022
10.1126/science.adf1017

COSMOCHEMISTRY

Nucleosynthetic isotope anomalies of zinc in meteorites constrain the origin of Earth's volatiles

Rayssa Martins*†, Sven Kuthning†‡, Barry J. Coles, Katharina Kreissig, Mark Rehkämper*

Material inherited from different nucleosynthesis sources imparts distinct isotopic signatures to meteorites and terrestrial planets. These nucleosynthetic isotope anomalies have been used to constrain the origins of material that formed Earth. However, anomalies have only been identified for elements with high condensation temperatures, leaving the origin of Earth's volatile elements unconstrained. We determined the isotope composition of the moderately volatile element zinc in 18 bulk meteorites and identified nucleosynthetic zinc isotope anomalies. Using a mass-balance model, we find that carbonaceous bodies, which likely formed beyond the orbit of Jupiter, delivered about half of Earth's zinc inventory. Combined with previous constraints obtained from studies of other elements, these results indicate that ~10% of Earth's mass was provided by carbonaceous material.

The Solar System formed from an initially hot protoplanetary disk; as the disk cooled, refractory (high condensation temperature) metals and metal oxides were the first materials to condense into solid dust grains, followed by Fe metal and Mg silicates (*1*). The relative abundances of these refractory elements in all Solar System bodies are similar to those found in the Sun and in Ivuna-type carbonaceous chondrite (CI) meteorites, which are thought to be representative of the Solar System as a whole (*2*). Volatile elements (which have condensation temperatures below those of Mg, Si, and Fe) condensed later, and all other groups of meteorites are depleted in volatiles as compared to CIs, to various extents (*2*–*5*). This depletion was inherited by the terrestrial planets; it has been measured for Earth, the Moon, Mars, and Mercury (*6*–*9*). It is unclear how these bodies obtained their volatile elements.

The origins of Earth's building blocks can be constrained using nucleosynthetic isotope anomalies measured between bulk meteorites and the bulk silicate Earth (BSE). These small isotopic variations are produced by heterogeneous distribution of isotopically distinct material during the formation of the Solar System. This distinct material is inherited from presolar processes, such as nucleosynthesis in stars or supernovae. Known nucleosynthetic anomalies reveal an isotopic dichotomy between carbonaceous and noncarbonaceous (NC) meteorites (*10*). The carbonaceous group consists primarily of carbonaceous chondrite (CC) meteorites, and the NC group consists of enstatite chondrite (EC) and ordinary chondrite (OC) meteorites. In comparison to NCs, CCs have higher abundances of neutron-rich isotopes of the Fe-peak elements Ca, Ti, Cr, and Ni, and lower abundances of the slow neutron capture (s-process) nuclides of Zr, Mo, Ru, and Pd (*10*–*15*). It has been proposed that this dichotomy resulted from spatial separation of two isotopic reservoirs, with NC material in the inner Solar System, and CC material on orbits outside that of Jupiter (*16*).

The isotopic signature of the BSE indicates that Earth accreted primarily from NC material, with a small but non-negligible contribution of CC material (*10*, *12*, *17*–*19*). However, previously reported nucleosynthetic isotope anomalies were for refractory metals, which have nearly uniform elemental abundance ratios throughout the Solar System. By contrast, ratios of volatile to refractory elements are variable in generally volatile-rich CCs and volatile-poor OCs, and even more so in ECs, with the latter presenting both severe volatile depletions and high volatile abundances that could potentially account for much of Earth's volatile inventory (*20*, *21*). Searches for nucleosynthetic isotope anomalies in the volatile elements K, Zn, and Te have been inconclusive (*22*–*26*).

Zinc isotope measurements

We measured bulk samples of 18 meteorites (table S1): 7 CCs from four major groups [CI, CM (Mighei type), CV (Vigarano type), and CO (Ornans type)], and 11 NC meteorites, including 7 OCs, 2 ECs, and 2 iron meteorites classified as members of the IAB complex. Zinc isotope data were measured using a multiple-collector inductively coupled plasma mass spectrometer (MC-ICP-MS) (*27*). Results are presented in ε^{i}Zn notation (eq. S1), denoting the parts-per-10,000 deviations of the $^{i}Zn/^{67}Zn$ ratio from the London Zn standard, which defines $\varepsilon^{i}Zn = 0$ for the BSE (*27*). Unless otherwise stated, all uncertainties are quoted as ±2 SE (twice the standard error of the mean).

We find small but resolvable deviations from the BSE for all meteorites that we measured (Fig. 1, Table 1, and table S2). The CCs have an excess in ε^{70}Zn (a neutron-rich isotope), a smaller excess for ε^{68}Zn, and a depletion in ε^{64}Zn. The NC meteorites have complementary patterns, with negative anomalies for ε^{70}Zn and ε^{68}Zn, but positive deviations for ε^{64}Zn. We find similar results if other isotope ratios are used for internal normalization [e.g., $^{64}Zn/^{67}Zn$ (fig. S1)]. Our measurements of BSE standards show no substantial deviations from $\varepsilon^{i}Zn = 0$ (Fig. 1).

Previous research has suggested that anomalies of the neutron-rich isotopes of other Fe-peak elements derive from either Type Ia supernovae (SNIa, thermonuclear explosions of white dwarfs) or electron-capture core collapse supernovae (ECSNe, explosions of certain types of massive star) (*11*, *28*–*30*). We therefore calculated the Zn isotope anomalies that would be produced by adding various mixtures of SNIa and ECSNe Zn to Earth's Zn (*27*), to assess whether the observed Zn isotope

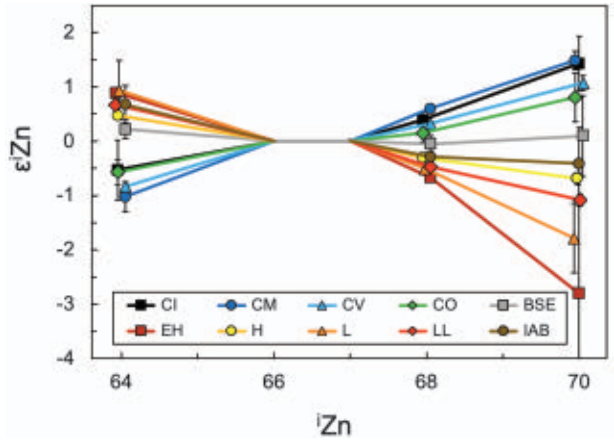


Fig. 1. Zinc isotope anomalies for each meteorite group and the BSE, in ^{i}Zn notation. Data are plotted for different groups of meteorites and the BSE (see the legend). Carbonaceous chondrites (CI, CM, CV, and CO chondrites) and noncarbonaceous-group meteorites [ECs, including two EHs (high iron); OCs, including H (high iron), L (low-iron), and LL (low-iron, low-metal) chondrites; and IAB complex irons] have complementary patterns with CCs enriched in ^{68}Zn and ^{70}Zn and depleted in ^{64}Zn relative to the BSE, which has $\varepsilon^{i}Zn = 0$. NCs have the opposite pattern, with depletions in ^{68}Zn and ^{70}Zn and enrichments in ^{64}Zn. There is no data for ε^{66}Zn and ε^{67}Zn because these isotopes are used for internal normalization. The meteorite and BSE data are listed in table S2; error bars indicate ±2 SE.

Department of Earth Science and Engineering, Imperial College London, London SW7 2AZ, UK.
*Corresponding author. Email: r.martins-pimentel19@imperial.ac.uk (R.M.); markrehk@imperial.ac.uk (M.R.)
†These authors contributed equally to this work.
‡Present address: Shimadzu Europa GmbH, 47269 Duisburg, Germany.

Table 1. Mean, minimum, and maximum ε64Zn, ε68Zn, and ε70Zn values of the CC and NC meteorite groups. Results of all groups are listed in table S1. Uncertainties are ±2 SE (twice the standard error of the mean).

	ε^{64}Zn	2 SE	ε^{68}Zn	2 SE	ε^{70}Zn	2 SE
			CC			
Mean	−0.81	0.16	0.31	0.12	1.20	0.20
Min	−1.22	0.15	0.15	0.16	0.98	0.52
Max	−0.53	0.55	0.65	0.41	1.61	0.83
			NC			
Mean	0.75	0.18	−0.45	0.09	−1.39	0.62
Min	0.44	0.003	−0.70	0.16	−3.93	1.68
Max	1.61	0.61	−0.21	0.16	0.06	0.05

anomalies could have been generated by these nucleosynthetic sources. Plotting ε^{70}Zn as a function of ε^{64}Zn (Fig. 2) shows that the meteorite data have a negative correlation, which matches our models (*27*). We therefore conclude that the measured Zn isotope anomalies have a nucleosynthetic origin.

Dichotomy between CC and NC meteorites

Comparing ε^{70}Zn to anomalies in ^{48}Ca, ^{50}Ti, and ^{54}Cr (fig. S2) confirms the dichotomy in compositions between CC and NC meteorites. However, most of our meteorite data do not follow the isotopic trends predicted by combining nucleosynthetic production models of the different elements (fig. S2). It is nevertheless plausible for both the ε^{70}Zn enrichments and the ^{48}Ca, ^{50}Ti, and ^{54}Cr enrichments to be produced in the same nucleosynthesis sites, because the differing trends could be produced by the different physical and chemical properties of the elements. For example, moderately volatile Zn has a much lower 50% condensation temperature (T_{C50} = 726 K) than Ca, Ti, and Cr [T_{C50} ≈ 1290 to 1570 K (*31*)], so condensation of isotopically distinct Zn from SNIa and/or ECSNe into presolar dust grains was likely less efficient than for the more refractory elements (*15*, *32*). Zinc also has more affinity for sulfide than Ca, Ti, and Cr do (*33*), so isotopically anomalous Zn could have been concentrated into distinct mineral phases, which could have been lost during subsequent thermal processing during Solar System formation (*34*). Although the presolar carrier phases of the Zn anomalies are unknown, Zn in CCs is predominantly hosted in the fine-grained meteorite matrix, with Zn elemental abundances that exceed those found in chondrules (silicate spherules embedded in meteorites) by factors of 3 to 6 (*35*–*37*). Leaching studies have shown that Zn in chondrites is primarily (>80%) hosted in sulfides and silicates (*36*, *38*).

Nucleosynthetic isotope anomalies of more refractory metals display substantial isotopic diversity within the CC and NC reservoirs (*39*), but the Zn isotope data are much less variable. The mean Zn isotope compositions of all CC groups that we measured overlap within their uncertainties; we performed analysis of variance (ANOVA) tests and found that the groups are statistically indistinguishable (*27*). Likewise, the ε^{64}Zn values of the ECs, OCs and IABs all overlap within their uncertainties (Fig. 1 and table S2), with ANOVA tests indicating that these groups are also statistically indistinguishable (*27*).

Mass-balance calculations

The complementary Zn isotope anomalies of CCs and NC meteorites indicate a mixed source for Zn in the BSE. Because Zn did not partition strongly into metal during Earth's core formation (table S3), the nucleosynthetic Zn isotope composition of the BSE was built up throughout Earth's accretion history (*40*). We therefore assume that the BSE and bulk Earth (BE) have identical Zn isotope compositions. Chondrites are often assumed to be the building blocks of planets because they (mostly) preserve their original elemental and isotopic compositions and can account for most of the Solar System's isotopic diversity (*17*, *41*). We take the overall mean values of CCs (ε^{64}Zn = −0.81 ± 0.16, n = 7, where n is the number of meteorites) and ECs and OCs [noncarbonaceous chondrites (NCCs)] (ε^{64}Zn = 0.76 ± 0.22, n = 9) as representative of the CC and NC material accreted by Earth (table S2). We then calculate a mass balance to infer the contributions of each to Earth's Zn:

$$\varepsilon^{64}\mathrm{Zn}_{\mathrm{BE}} = x^{\mathrm{Zn}}_{\mathrm{CC}} \cdot \varepsilon^{64}\mathrm{Zn}_{\mathrm{CC}} + x^{\mathrm{Zn}}_{\mathrm{NCC}} \cdot \varepsilon^{64}\mathrm{Zn}_{\mathrm{NCC}} \quad (1)$$

where x^{Zn} is the Zn fraction derived from each reservoir, and $x^{\mathrm{Zn}}_{\mathrm{CC}} + x^{\mathrm{Zn}}_{\mathrm{NCC}} \equiv 1$. We performed a Monte Carlo simulation with 10,000 trials (*27*), in which the ε^{64}Zn compositions of the CC and NCC sources were allowed to vary within ±2 SE of their respective mean values, the fractions of Zn from each source ranged between 0 and 1, and the BE composition of $\varepsilon^{64}\mathrm{Zn}_{\mathrm{BE}}$ = 0 ± 0.17 was assigned an uncertainty corresponding to the long-term reproducibility of four BSE standards (table S2).

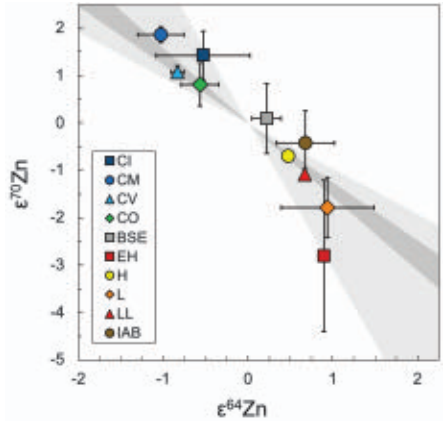


Fig. 2. ε^{70}Zn as a function of ε^{64}Zn for meteorite groups and the BSE. Data points indicate meteorite groups and BSE standards, using the same colors and symbols as Fig. 1. Shaded areas indicate predictions from models of Zn nucleosynthesis in Type Ia supernovae (*30*) in light gray, and electron capture supernovae (*28*) in dark gray (*27*). The meteorite and BSE data are listed in table S2; error bars indicate ±2 SE.

To reproduce the BE Zn isotope composition, we find that 48 ± 15% (uncertainty is 2 SD, twice the standard deviation) of Earth's Zn was derived from the CC reservoir, and the remainder from NCCs (fig. S3). In other words, CCs and NCCs contributed about equally to Earth's Zn. We obtain consistent (but more uncertain) results if other εZn values (e.g., ε^{70}Zn) or individual meteorite group means are used in the mass-balance equation (table S10). We adopt the results from ε^{64}Zn because its isotope anomalies are larger and more precise than those of other zinc isotopes.

We extended our mass-balance equation to constrain the mass fraction of carbonaceous material of all elements (x_{CC}) that was accreted by Earth:

$$\varepsilon^{64}\mathrm{Zn}_{\mathrm{BE}} = \frac{x_{\mathrm{CC}} \cdot [\mathrm{Zn}]_{\mathrm{CC}} \cdot \varepsilon^{64}\mathrm{Zn}_{\mathrm{CC}} + x_{\mathrm{NCC}} \cdot [\mathrm{Zn}]_{\mathrm{NCC}} \cdot \varepsilon^{64}\mathrm{Zn}_{\mathrm{NCC}}}{x_{\mathrm{CC}} \cdot [\mathrm{Zn}]_{\mathrm{CC}} + x_{\mathrm{NCC}} \cdot [\mathrm{Zn}]_{\mathrm{NCC}}} \quad (2)$$

and

$$[\mathrm{Zn}]_{\mathrm{BE}} = x_{\mathrm{CC}} \cdot [\mathrm{Zn}]_{\mathrm{CC}} + x_{\mathrm{NCC}} \cdot [\mathrm{Zn}]_{\mathrm{NCC}} \quad (3)$$

where x is the mass fraction of material from each reservoir, $x_{\mathrm{CC}} + x_{NCC} \equiv 1$, and [Zn] denotes Zn elemental abundances. We performed further Monte Carlo simulations of these equations using constraints on the Zn elemental abundance of the BE and CC material (*27*). Although the Zn elemental abundance of the BSE is well defined [54 ± 2 μg g^{-1} (table S3)], $[\mathrm{Zn}]_{\mathrm{BE}}$ is less well constrained at 40 to 70 μg g^{-1}, because the extent of Zn partitioning into

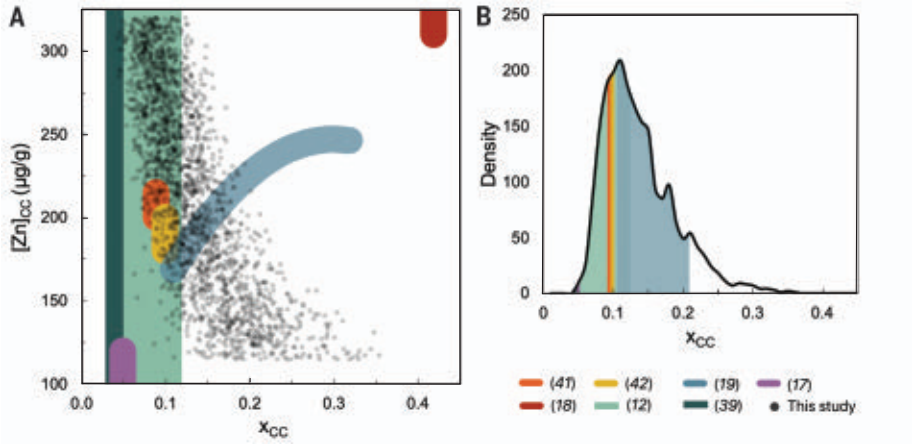


Fig. 3. Estimates for the mass fractions of accreted CC material (x_{CC}). (**A**) Zn elemental abundances plotted as a function of x_{CC}. Black dots show individual results from our Monte Carlo simulations, and colored regions indicate x_{CC} estimates from previous studies in orange (*41*), yellow (*42*), light blue (*19*), purple (*17*), red (*18*), light green (*12*), and dark green (*39*). (**B**) Probability density (black curve) of the estimated x_{CC} from our Monte Carlo simulations. Shaded regions indicate estimates from previous studies, using the same colors shown in (A). Our simulations have the highest density at $x_{CC} \approx 11\%$, equivalent to mean $[Zn]_{CC}$ concentrations of ~150 to 320 μg g^{-1}. Considering additional constraints from studies of other elements (*12*, *19*, *41*, *42*), the mean $[Zn]_{CC}$ elemental abundance is limited to ~170 to 220 μg g^{-1}, compatible with CM chondrites (table S3).

Earth's core is uncertain (table S3). For the CC material, the Zn elemental abundance was allowed to vary between 115 and 320 μg g^{-1}, because this variance encompasses the mean Zn abundances of CI, CM, and CO-CV chondrites (table S3). The Zn elemental abundances of ECs and OCs span a much wider range [7.5 to 519 μg g^{-1} (table S3)], so we treat $[Zn]_{NCC}$ as a free parameter. We find that 13 ± 10% (2 SD) of Earth's mass is derived from CC material (Fig. 3). At the lower end of this range (at $x_{CC} \approx$ 3 to 5%), Earth must have incorporated CI-like material; larger x_{CC} mass fractions require that increasing proportions of more volatile-depleted CCs were accreted (Fig. 3).

Comparison to other elements

We compare these results to previous studies that applied refractory-element and O isotope compositions to evaluate the origin of Earth's accreted material (Fig. 3). A CC mass fraction of 42% CI material has previously been proposed (*18*), based on differences in ε^{48}Ca between the BSE and other inner Solar System bodies. For such a high CC fraction to be consistent with our Zn results, the BE Zn elemental abundance would need to be ~280 μg g^{-1} to produce $\varepsilon Zn_{BE} \approx 0$, far higher than the estimated BE Zn abundance [40 to 70 μg g^{-1} (table S3)]. Another study (*17*) inferred a much smaller CC contribution, of 5% CO-CV material, based on an inversion of refractory-element and O isotope compositions. In that case, the Zn contribution from the CO-CV component is too small to be compatible with the Zn isotope mass balance, requiring the BE to have only ~12 μg g^{-1} Zn, much lower than estimated (table S3).

Studies based on O isotope compositions have proposed that Earth accreted from a mix of ECs and OCs, with addition of 4% CI plus 5% CV material (*41*) or 10% CM-like material (*42*); these are consistent with our results for Zn. It has also been proposed (*19*) that Earth formed primarily from noncarbonaceous material composed of OCs and angrites (noncarbonaceous achondritic meteorites), with addition of 3 to 21% CI and 8 to 11% CV material. Although this estimate is consistent with our Zn isotope mass balance (Fig. 3), and we do not detect substantial variability within the NC reservoir, the lack of Zn isotope data for angrite meteorites prevents us from drawing firm conclusions about this scenario.

It is difficult to find a combination of meteoritic materials that can explain the BSE Mo isotope composition (*39*). It has been proposed that additional NC component, which was present in the early Solar System but is not represented in measured meteorites, is enriched in both s-process isotopes and neutron-rich nuclei of Fe-peak elements; if such a reservoir exists, it implies the accretion of ~4% CC material (*39*). This scenario can be reconciled with our Zn isotope data if Earth accreted CI-like material, but it conflicts with O isotope constraints (*41*, *42*). If the unidentified component is instead carbonaceous (*12*), the mass balance would allow for a larger CC contribution that is more consistent with our Zn isotope data. In the latter scenario, 30 to 60% of the BSE Mo budget would be derived from the carbonaceous reservoir (*12*). This can be reconciled with our Zn isotope constraints because Mo is more easily incorporated into Earth's core than Zn, so the BSE Mo isotope composition reflects only the final 10 to 20% of accretion, implying a minimum Mo-based x_{CC} of 3 to 12%. Therefore, a ~10% CC contribution composed of CMs or a mixture of CI, CM, and CO-CV chondrites can be reconciled with refractory-element nucleosynthetic isotope anomalies (*12*, *19*), O isotope constraints (*41*, *42*), and our Zn results (Fig. 3).

Origin of Earth's zinc

Our mass-balance calculations assume that the CC material and Earth did not experience Zn loss before or during accretion; the results then suggest that the NC material was, on average, depleted in Zn relative to CCs. Although some NC meteorites have high Zn elemental abundances [>300 μg g^{-1} (table S3)], an accretion scenario with a volatile-rich NC component and volatile-depleted CC material can only be accommodated in the Zn mass balance if x_{CC} is large. However, that is difficult to reconcile with refractory-element isotope compositions, which indicate a smaller contribution from CCs. Substantial loss of Zn from planets during their accretion, or directly from the accreting material, would introduce mass-dependent isotope fractionation. This would leave planets and meteorites enriched in heavy isotopes of volatile elements, as measured for Zn and K in lunar rocks (*43*, *44*). A similar signature is not observed for the BSE, indicating that Earth's volatile depletion arose from processes in the protoplanetary disk rather than from evaporation during subsequent accretion (*43*, *44*).

If the volatile elements in the BSE are representative of the accreting material, our results also constrain the mean Zn abundance of the NC material added to Earth. If about half of Earth's Zn, with $[Zn]_{BE}$ = 40 to 70 μg g^{-1}, was provided by NC bodies with a combined mass of 77 to 97% of Earth's, this implies a Zn abundance of 18 to 48 μg g^{-1} in the NC material. That is consistent with measured ranges for ECs (7.5 to 519 μg g^{-1}) and OCs [14 to 158 μg g^{-1} (table S3)]. Elements with condensation temperatures similar to Zn (500 K < T_{C50} < 750 K) are depleted by similar proportions in CCs, with an equivalent trend also observed for the BSE (*3*).

In conclusion, our Zn isotope data indicate that about half of Earth's Zn inventory was derived from carbonaceous material, which provided ~10% of the planet's mass. It is therefore likely that carbonaceous material also made substantial contributions to Earth's abundances of other moderately and highly volatile elements.

REFERENCES AND NOTES

1. K. Lodders, *Astrophys. J.* **591**, 1220–1247 (2003).
2. J. T. Wasson, G. W. Kallemeyn, *Soc. Lond. A.* **325**, 535–544 (1988).
3. N. Braukmüller, F. Wombacher, C. Funk, C. Münker, *Nat. Geosci.* **12**, 564–568 (2019).
4. M.-S. Wang, M. E. Lipschutz, *Geochim. Cosmochim. Acta* **71**, 1062–1073 (2007).

5. M.-S. Wang, M. E. Lipschutz, *Environ. Chem.* **2**, 215–226 (2005).
6. C. Allègre, G. Manhès, É. Lewin, *Earth Planet. Sci. Lett.* **185**, 49–69 (2001).
7. H. S. C. O'Neill, *Geochim. Cosmochim. Acta* **55**, 1159–1172 (1991).
8. G. J. Taylor, *Chem. Erde* **73**, 401–420 (2013).
9. P. N. Peplowski *et al*., *Science* **333**, 1850–1852 (2011).
10. P. H. Warren, *Earth Planet. Sci. Lett.* **311**, 93–100 (2011).
11. N. Dauphas *et al*., *Earth Planet. Sci. Lett.* **407**, 96–108 (2014).
12. G. Budde, C. Burkhardt, T. Kleine, *Nat. Astron.* **3**, 736–741 (2019).
13. M. Fischer-Gödde, C. Burkhardt, T. S. Kruijer, T. Kleine, *Geochim. Cosmochim. Acta* **168**, 151–171 (2015).
14. W. Akram, M. Schönbächler, S. Bisterzo, R. Gallino, *Geochim. Cosmochim. Acta* **165**, 484–500 (2015).
15. M. Ek, A. C. Hunt, M. Lugaro, M. Schönbächler, *Nat. Astron.* **4**, 273–281 (2020).
16. T. S. Kruijer, C. Burkhardt, G. Budde, T. Kleine, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6712–6716 (2017).
17. N. Dauphas, *Nature* **541**, 521–524 (2017).
18. M. Schiller, M. Bizzarro, V. A. Fernandes, *Nature* **555**, 507–510 (2018).
19. C. Fitoussi, B. Bourdon, X. Y. Wang, *Earth Planet. Sci. Lett.* **434**, 151–160 (2016).
20. K. Lodders, B. Fegley, *The Planetary Scientist's Companion* (Oxford Univ. Press, 1998).
21. L. Piani *et al*., *Science* **369**, 1110–1113 (2020).
22. Y. Ku, S. B. Jacobsen, *Sci. Adv.* **6**, eabd0511 (2020).
23. N. X. Nie *et al*., *Sci. Adv.* **7**, eabl3929 (2021).
24. F. Moynier, N. Dauphas, F. A. Podosek, *Astrophys. J.* **700**, L92–L95 (2009).
25. F. Moynier, D. Vance, T. Fujii, P. Savage, *Rev. Mineral. Geochem.* **82**, 543–600 (2017).
26. M. A. Fehr *et al*., *Geochim. Cosmochim. Acta* **69**, 5099–5112 (2005).
27. Materials and methods are available as supplementary materials.
28. S. E. Woosley, *Astrophys. J.* **476**, 801–810 (1997).
29. L. P. Qin, C. M. O. Alexander, R. W. Carlson, M. F. Horan, T. Yokoyama, *Geochim. Cosmochim. Acta* **74**, 1122–1145 (2010).
30. S. Wanajo, H.-T. Janka, B. Müller, *Astrophys. J. Lett.* **767**, L26 (2013).
31. B. J. Wood, D. J. Smythe, T. Harrison, *Am. Mineral.* **104**, 844–856 (2019).
32. K. K. Marhas, P. Sharda, *Astrophys. J.* **853**, 12 (2018).
33. A. G. Christy, *Eur. J. Mineral.* **30**, 193–204 (2018).
34. A. Trinquier *et al*., *Science* **324**, 374–376 (2009).
35. P. A. Bland *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13755–13760 (2005).
36. E. A. Pringle, F. Moynier, P. Beck, R. Paniello, D. C. Hezel, *Earth Planet. Sci. Lett.* **468**, 62–71 (2017).
37. J. L. Hellmann, T. Hopp, C. Burkhardt, T. Kleine, *Earth Planet. Sci. Lett.* **549**, 116508 (2020).
38. J.-M. Luck, D. B. Othman, F. Albarède, *Geochim. Cosmochim. Acta* **69**, 5351–5363 (2005).
39. C. Burkhardt *et al*., *Sci. Adv.* **7**, eabj7601 (2021).
40. B. Mahan *et al*., *Geochim. Cosmochim. Acta* **235**, 21–40 (2018).
41. K. Lodders, *Space Sci. Rev.* **92**, 341–354 (2000).
42. F. Albarède *et al*., *Icarus* **222**, 44–52 (2013).
43. P. A. Sossi, O. Nebel, H. S. C. O'Neill, F. Moynier, *Chem. Geol.* **477**, 73–84 (2018).
44. Z. Tian *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2101155118 (2021).

ACKNOWLEDGMENTS

We are grateful to the NASA Johnson Space Center (K. Righter, C. Satterwhite), the Natural History Museum London (N. Almeida, C. Smith), the Field Museum of Natural History (P. R. Heck, J. L. Holstein), the Smithsonian National Museum of Natural History (T. McKoy, J. E. Post), the Senckenberg Gesellschaft für Naturforschung (J. Zipfel) for providing meteorite samples. We thank C. Hémond, S. Mukasa, and K. Crocket for contributing the Earth rocks and N. Dauphas for assistance with the nucleosynthetic modeling. We thank F. Wombacher, C. Münker, T. van de Fliert, and also M. Fischer-Gödde and two other anonymous reviewers for constructive comments that helped improve the manuscript. **Funding:** R.M. was supported by an Imperial College London President's PhD Scholarship. S.K. was supported by the Janet Watson PhD Scholarship of the Department of Earth Science and Engineering of ICL. M.R. was funded by UKRI STFC grant ST/W001179/1. **Author contributions:** Conceptualization and funding: M.R. Method development and data acquisition: S.K., R.M., B.J.C., K.K., and M.R. Data Analysis: S.K., R.M., and M.R. Interpretation and modelling: R.M. and M.R. Writing: R.M., M.R., and S.K. (analytical methods).

Competing interests: The authors declare no competing interests. **Data and materials availability:** The meteorite and BSE samples were obtained from the sources listed in table S1 and are described in the supplementary materials. Our mean Zn isotope data for each meteorite and BSE sample are listen in table S2. The results of the mass-balance calculations are provided in tables S10 and S11.

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn1021
Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S11
References (*45*–*84*)
Data S1

Submitted 3 November 2021; accepted 22 December 2022
10.1126/science.abn1021

COSMOCHEMISTRY

Meteorites have inherited nucleosynthetic anomalies of potassium-40 produced in supernovae

Nicole X. Nie[1]*†, Da Wang[1,2]*†, Zachary A. Torrano[1], Richard W. Carlson[1], Conel M. O'D. Alexander[1], Anat Shahar[1]

Meteorites record processes that occurred before and during the formation of the Solar System in the form of nucleosynthetic anomalies: isotopic compositions that differ from the Solar System patterns. Nucleosynthetic anomalies are rarely seen in volatile elements such as potassium at bulk meteorite scale. We measured potassium isotope ratios in 32 meteorites and identified nucleosynthetic anomalies in the isotope potassium-40. The anomalies are larger and more variable in carbonaceous chondrite (CC) meteorites than in noncarbonaceous (NC) meteorites, indicating that CCs inherited more material produced in supernova nucleosynthesis. The potassium-40 anomaly of Earth is close to that of the NCs, implying that Earth's potassium was mostly delivered by NCs.

Nucleosynthetic anomalies are isotopic abundance patterns that deviate from the apparently uniform Solar System patterns. Their presence in primitive meteorites provides information about the production of nuclides in stars (stellar nucleosynthesis), the incorporation of elements into the Solar System during its formation, and the subsequent processing and transport of material in the solar nebula (the cloud of gas and dust from which the Solar System formed by condensation). Nucleosynthetic anomalies have been identified in bulk meteorites for several refractory elements, i.e., those that condense from the solar nebula before or concurrent with the major elements Mg, Si, and Fe (*1*), namely Ca (*2*), Ti (*3*), Cr (*4*), Fe (*5*), Ni (*6*), Mo (*7*), and Ru (*8*, *9*). However, despite sensitive searches (*10*), these anomalies are rarely seen in volatile elements, i.e., those that condense after the major elements.

The lack of detectable nucleosynthetic anomalies in volatile elements has been interpreted as indicating that extensive thermal processing in the hot solar nebula volatilized solid dust particles and homogenized the distribution of volatile elements in the Solar System, but not the refractory elements. For the moderately volatile element (MVE) potassium (K), variations of the $^{41}K/^{39}K$ ratios measured in bulk meteorites might reflect nucleosynthetic anomalies (*11*). However, the $^{41}K/^{39}K$ ratios correlate with $^{87}Rb/^{85}Rb$ ratios in carbonaceous meteorites, which could arise if K and Rb experienced mass-dependent isotope fractionation (processes that produce different isotope abundances in proportion to their masses) when they were evaporated and recondensed into chondrules (millimeter-sized glassy spherules) during meteorite formation (*12*). These competing interpretations cannot be distinguished with measurements of only two isotopes (^{41}K and ^{39}K), because the same ratio could arise from either mass-dependent fractionation or nucleosynthetic anomalies.

If the variations in the $^{41}K/^{39}K$ ratio are due to nucleosynthetic anomalies, then measurements of a third isotope (such as ^{40}K) will deviate from the mass-dependent relationship defined by the other two isotopes. However, ^{40}K has not been measured in meteorites because of its much lower natural abundance (the mole fraction of ^{40}K is 0.012% of total K atoms). Previous measurements of K isotopes (*11*, *12*) were limited by instrumental interferences from ^{40}Ar and ^{40}Ca, which are far more abundant than ^{40}K.

K isotope measurements

We sought to overcome this limitation using thermal ionization mass spectrometry (TIMS)

[1]Earth and Planets Laboratory, Carnegie Institution for Science, Washington, DC 20015, USA. [2]International Center for Planetary Science, College of Earth Sciences, Chengdu University of Technology, 610059 Chengdu, China.
***Corresponding author. Email: nnie@carnegiescience.edu (N.X.N.); dwang@carnegiescience.edu (D.W.)**
†These authors contributed equally to this work.

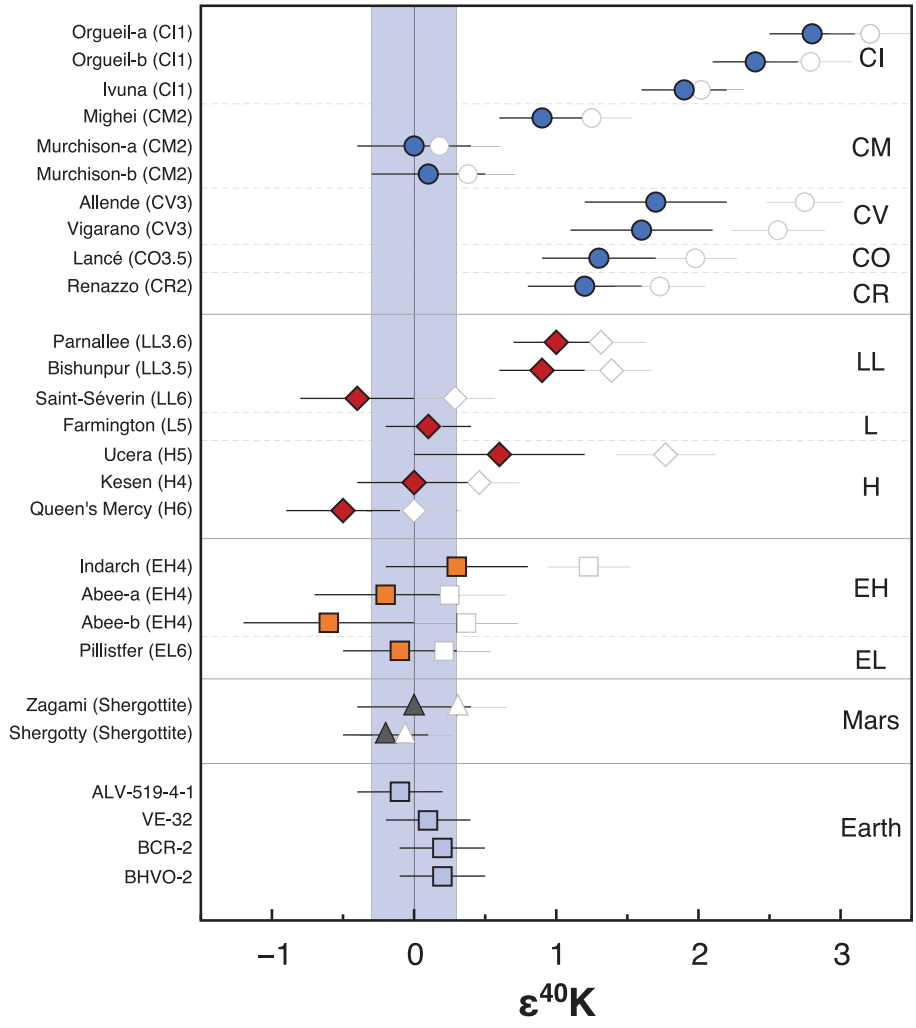


Fig. 1. K isotopic anomalies of meteorites with low CRE corrections. For each meteorite or terrestrial standard (labeled on the left), the empty and colored symbols represent K isotopic anomalies before and after CRE correction, respectively. Each color and symbol type indicates a class of meteorites (labeled on the right) that are separated by solid gray horizontal lines; dashed gray horizontal lines separate subclasses. The blue shaded region indicates the terrestrial range. Error bars indicate 95% confidence interval. Only samples with CRE corrections of $\varepsilon^{40}K \leq 1$ are shown here; fig. S3 shows all of the samples. The analytical uncertainty is ~±0.4 on this scale.

to measure all three isotopes of K without interference from other elements. Meteorites are broadly divided into carbonaceous chondrite (CC) and noncarbonaceous (NC) families on the basis of their physical, chemical, and isotopic properties. We measured 32 bulk meteorites (table S1 and data S1A): 10 CCs (from the CI, CM, CV, CO, and CR subgroups), 18 NCs [enstatite chondrite (EC) groups EH and EL and ordinary chondrite (OC) groups H, L, and LL], and four martian meteorites. Because K is highly soluble in water, all but one of the samples that we selected were collected soon after observed falls to Earth to minimize any potential terrestrial contamination introduced by rain or dew. The K isotopic data are expressed using the $\varepsilon^{40}K$ notation as follows:

$$\varepsilon^{40}K \equiv 10,000\left[\frac{\left({}^{40}K/{}^{39}K\right)^{*}_{\text{sample}}}{\left({}^{40}K/{}^{39}K\right)^{*}_{\text{terrestrial standard}}} - 1\right] \quad (1)$$

where asterisks indicate that corrections have been applied to remove the mass-dependent fractionation component. The $\varepsilon^{40}K$ value would be zero for mass-dependent isotope fractionation and nonzero for isotopic anomalies.

Our measured $\varepsilon^{40}K$ variations in the bulk meteorite samples range from ~0 to +4.5, with typical uncertainties of ±0.4 (table S1). These variations are similar to the ranges of nucleosynthetic anomalies previously found for refractory elements (*2*–*9*).

Origins of the isotopic anomaly

These deviations of K isotopes from the terrestrial standards could have four causes: (i) inaccuracy of our correction for mass-dependent fractionation leading to an apparent anomaly; (ii) radioactive decay of short-lived ^{41}Ca to ^{41}K; (iii) excess ^{40}K through the reaction $^{40}Ca(n, p)\,^{40}K$ or spallation of Fe-group nuclides produced by cosmic ray exposure (CRE) of the meteorites during transfer to Earth (*13*); or (iv) pre–Solar System anomalies inherited from stellar nucleosynthesis.

To test scenario (i), we calculated corrections using several different mass dependencies in addition to the fiducial exponential function (*14*). We found a maximum apparent $\varepsilon^{40}K$ of ±0.06 (fig. S1), which is ~10 times smaller than the analytical uncertainty, so scenario (i) cannot be the cause of the observed variation. In scenario (ii), radioactive decay of ^{41}Ca would produce an excess in ^{41}K, and therefore a deficit in ^{40}K after correction for mass-dependent fractionation. This is inconsistent with our measurements of zero or positive $\varepsilon^{40}K$ values. In addition, the low initial $^{41}Ca/^{40}Ca$ ratio of the Solar System [~10^{-9} to 10^{-8} (*15*)] would produce only enough ^{41}K to contribute −0.02 to $\varepsilon^{40}K$ in bulk samples (*14*), which is 10 times smaller than our uncertainties and thus would not be detectable. For scenario (iii), we calculated the CRE-induced ^{40}K excess ($\varepsilon^{40}K_{CRE}$; see table S1 and fig. S2) using previously constrained ^{40}K production rates (*13*); the CRE age of each meteorite; and their K, Ca, and Fe concentrations (*14*). We found that $\varepsilon^{40}K_{CRE}$ is generally small in CCs, but larger in NCs because of their longer CRE ages. After deducting $\varepsilon^{40}K_{CRE}$ from the measured $\varepsilon^{40}K$, many meteorites still show ^{40}K excesses ($\varepsilon^{40}K_{Corr}$; see table S1 and fig. S3). To further determine whether the elevated $\varepsilon^{40}K_{Corr}$ values are due to nucleosynthetic anomalies, we focused on the samples with the lowest CRE corrections. About half of our

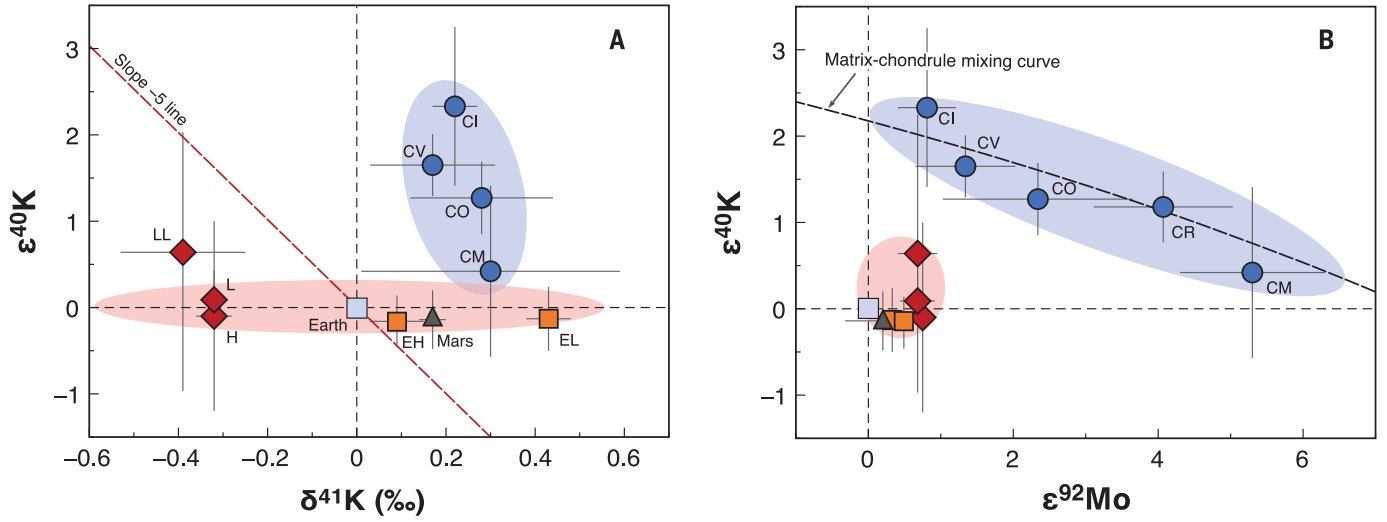


Fig. 2. Comparison of ^{40}K anomalies, δ^{41} values, and ^{92}Mo anomalies for different meteorite groups. Colored symbols represent different groups. Error bars indicate 95% confidence interval. Blue and pink areas represent CC and NC families, respectively. The horizontal and vertical dashed lines indicate the Earth values. (**A**) Average ε^{40}K for each meteorite class (after CRE correction; data S1D) shown as a function of δ^{41}K (*11*). The red dashed line indicates the slope −5 trend that would be expected if ^{41}K were anomalous and not ^{40}K. The data are not consistent with that interpretation. (**B**) Average ε^{40}K for each meteorite class as a function of ε^{92}Mo nucleosynthetic anomalies (data S1, D and E). The data are consistent with a CC-NC dichotomy. The dashed black curve shows a calculated matrix-chondrule mixing curve for CCs.

samples have small differences between the measured ε^{40}K and the values corrected for CRE (fig. S4). After this reduction in samples, most CCs show ^{40}K excesses (Fig. 1), whereas all but two NCs have $\varepsilon^{40}K_{Corr}$ values that are consistent with the terrestrial value. We infer that the observed ^{40}K excesses cannot be explained solely by CRE, and they must reflect pre–Solar System processes. We regard stellar nucleosynthesis as the most likely cause of the K isotopic anomalies.

With only three isotopes, we cannot determine from the measurements alone which isotope is anomalous. Expressing the results as ε^{40}K implicitly assumes that ^{40}K is anomalous, not the other two isotopes. However, a previous study (*11*) suggested that the variation of the ^{41}K/^{39}K ratio (without correction for mass-dependent fractionation) between meteorite groups indicates anomalies in ^{41}K. If this is correct, meaning that ^{41}K and not ^{40}K is the anomalous isotope, then the ^{41}K anomaly would have been reassigned to ^{40}K during our mass fractionation correction. In that case, our ε^{40}K values would correlate with their δ^{41}K values

$$\delta^{41}K \equiv 1000\left[\frac{\left(^{41}K/^{39}K\right)_{sample}}{\left(^{41}K/^{39}K\right)_{terrestrial\ standard}} - 1\right] \quad (2)$$

with a slope of ~−5 (because $-\frac{10{,}000\left[M\left(^{40}K\right)-M\left(^{39}K\right)\right]}{1000\left[M\left(^{41}K\right)-M\left(^{39}K\right)\right]} \simeq -5$, where M is the mass of each isotope). However, we did not find such a correlation (Fig. 2A), indicating that the anomaly is in ^{40}K not ^{41}K. This is consistent with a previous suggestion (*12*) that the δ^{41}K variations among the CC groups reflect mass-dependent fractionation associated with chondrule formation and thus are not of nucleosynthetic origin.

Comparison with nucleosynthesis models

Models of stellar nucleosynthesis predict that ^{40}K is most likely the anomalous isotope. All three K isotopes are primarily produced during core-collapse explosions of massive stars (type II supernovae), but ^{40}K is produced by more nuclear reaction pathways than the other two isotopes (*16*, *17*). Using a set of existing supernova nucleosynthesis models (*18*), we tracked the production of K in a 25 solar mass ($M_\odot$) supernova (*14*). We assumed that the yield of this model is representative of the average from type II supernovae of all masses (*14*). The results show that ^{40}K is produced at different stages of the supernova than ^{39}K and ^{41}K (fig. S5). The production of ^{39}K and ^{41}K (including radiogenic ^{39}K and ^{41}K from short-lived ^{39}Ar and ^{41}Ca, respectively) occurs mainly in the O/Si/S zone of the supernova due to fusion of oxygen (O burning). By contrast, O burning accounts for <20% of ^{40}K production. Most ^{40}K is instead produced by C burning and the weak slow neutron capture process (weak s-process) in other zones (*14*). Individual zones of supernovae are usually homogeneously mixed because of convection, but different zones vary in composition (*18*). Therefore, the synthesis of ^{40}K is temporally and spatially decoupled from that of ^{39}K and ^{41}K, so ^{40}K is more likely to vary in abundance in different parts of the supernova ejecta.

Alternative nucleosynthesis origins could produce anomalies in ^{41}K due to radioactive decay of ^{41}Ca. This might occur in asymptotic giant branch (AGB) stars, where little K is synthesized, but ^{41}Ca could be produced through neutron capture by ^{40}Ca (*19*). Alternatively, cooling and condensation of supernova ejecta could fractionate Ca from K, because Ca condenses at a higher temperature than K (*16*). If condensation occurs before the decay of ^{41}Ca [half-life of 9.94×10^4 years (*20*)], then the first dust grains to condense would have high Ca/K ratios, leading to radiogenic ^{41}K excesses once the ^{41}Ca decays. A complementary low Ca/K reservoir with ^{41}K depletion would be expected in the grains that condensed later, but the degree of ^{41}K depletion might be too small to be detectable. ^{41}K anomalies have only been detected as excesses in microscale presolar grains (*21*). Any excesses in ^{41}K would appear as deficits in ^{40}K in our measurements, as described above, so if any ^{41}K excesses are present, then our measured ε^{40}K values are lower limits.

We would expect large-scale ^{40}K heterogeneity in the interstellar medium. The abundance of ^{40}K decreases with time by decaying to ^{40}Ca and ^{40}Ar, with a half-life of 1.25 billion years (*22*). However, multiple generations of supernovae would inject freshly synthesized ^{40}K into their local interstellar medium, causing an enrichment of ^{40}K in those areas compared with the average interstellar medium. It is not known from which area the Solar System K is inherited, but if we assume that the difference between the ^{40}K enrichment in the supernova nucleosynthesis model and that in the Solar System reflects the decay of ^{40}K, then we can calculate an average delay of ~3 billion years between K production during

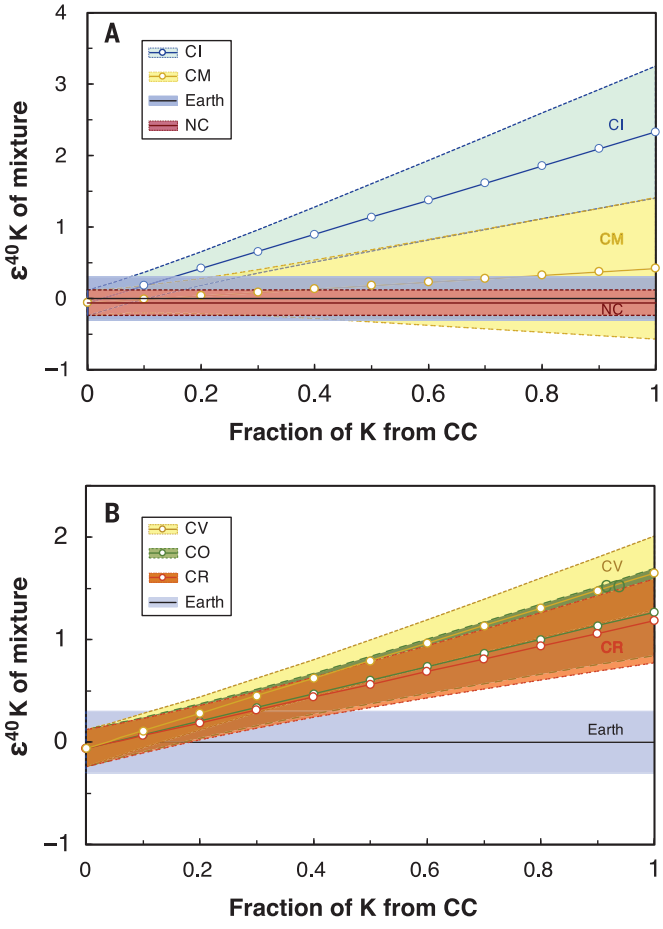


Fig. 3. Calculated fractions of Earth's K that could have originated from CC and NC reservoirs. The black line (with blue uncertainty region) shows Earth's ε^{40}K value of 0.0 ± 0.3, constrained by terrestrial standards. The results are split between two panels (A) and (B) with different vertical scales for clarity. For each CC meteorite group, we calculated the ε^{40}K that would result from a mixture between the NC reservoir (ε^{40}K = −0.06 ± 0.18, indicated by the red line and shaded uncertainty region) and different fractions of CC meteorites. Each CC group is represented by a solid line and a shaded uncertainty region in the same color (see legend); white dots denote the CC fraction from 0 to 1 in increments of 0.1. (**A**) Because the CM group has ε^{40}K = 0.42 ± 0.99, which overlaps with that of the Earth, any fraction of K from CMs is allowed. However, a large contribution (>20%) from CMs is inconsistent with other isotopic systems (*33*). If CI is the CC component, then the amount of K from CI is restricted to ≲10%. (**B**) For CV, CO, and CR chondrites, the K fraction is restricted to ≲20%. These calculations assume no volatile loss during Earth's accretion.

supernova nucleosynthesis and the birth of the Solar System (*14*).

Dichotomy between CCs and NCs

Nucleosynthetic anomalies of refractory elements show distinct patterns for CCs and NCs (*23*, *24*). This CC-NC dichotomy has been attributed to barriers separating the inner Solar System (where NCs formed) from the outer Solar System (where CCs formed), combined with infall of additional material from the surrounding molecular cloud with a composition varying over time (*25*, *26*). The isotopic anomalies that we found for K are consistent with the CC-NC dichotomy previously defined by anomalies in refractory elements (Fig. 2B and fig. S6). The NCs in our sample have ε^{40}K values close to the terrestrial value, whereas the CCs show a wide range of excesses, suggesting that the CC reservoir received higher contributions of material produced in type II supernovae. However, there is no correlation between the anomalies of K and those of refractory elements, except for Mo in the CCs, which correlates with K [Fig. 2B and fig. S7 (*14*)]. The lack of correlation with most refractory nuclides is probably because these nuclides are produced by different nucleosynthetic processes [many of the relevant nuclides are primarily produced in type Ia supernovae, e.g., ^{48}Ca, ^{50}Ti, ^{54}Cr, and ^{64}Ni (*17*)] or because they condense onto different dust grains at different times than K because they are more refractory. In contrast to K isotopes, which are mainly synthesized in supernovae, bulk meteorites show Mo isotopic anomalies consistent with an s-process deficit, suggesting that Mo is predominantly produced in AGB stars (*27*–*29*). Therefore, the K and Mo anomalies are unlikely to have the same nucleosynthetic origin.

The correlation between K and Mo could reflect volatility control, because K is volatile and Mo becomes volatile under oxidizing conditions (*30*). The anomalies of refractory elements among the CC groups might reflect mixing between the three basic chondritic components (chondrules, refractory inclusions, and fine-grained matrix with similar composition to CI chondrites) in varying proportions (*31*). The abundance of refractory inclusions is related to the magnitude of the anomalies in bulk CCs for some refractory elements (e.g., Ca and Ti) but not for Mo (*31*). Mass balance considerations show that chondrules and matrix are the two dominant reservoirs for Mo (*29*) and K (*12*). We calculated matrix-chondrule mixing for K and Mo anomalies (*14*), finding that the CCs follow the theoretical mixing curves (Fig. 2B and fig. S7) between two end members: (i) matrix, with a large ^{40}K excess and ^{92}Mo deficit, and (ii) chondrules, with small or opposite anomalies. We expect the matrix to preserve the largest anomalies because it is less thermally processed than chondrules. The anomalies in matrix are consistent with nucleosynthesis in supernovae producing ^{40}K excesses and that in AGB stars producing ^{92}Mo deficits [figs. S5, S8, and S9 (*14*)]. Chondrules were formed by high-temperature heating of CI-like matrix (or dust with similar composition), which could have preferentially destroyed carriers of Mo and K anomalies, leading to smaller or opposite anomalies. Subsequent mixing between matrix and chondrules, in variable proportions, could have produced the K-Mo trend among CC groups (Fig. 2B). However, if this mixing model is correct, then the magnitudes of both K and Mo anomalies should increase with increasing mass fractions of matrix in the CC groups. We found that this is only true if the CM group is excluded (fig. S10). It is unclear why the CM group might be an outlier; possibilities include sampling bias or that CMs incorporated material from different reservoirs than other CC groups (*14*).

Origin of Earth's K

The implications of our ε^{40}K measurements for the sources of Earth's K and potentially other MVEs (*14*) are as follows. For unknown reasons, Earth is strongly depleted in MVEs compared with CI chondrites, which are thought to reflect the bulk Solar System composition. One possibility is that Earth initially accreted with a low abundance of volatile elements, then acquired them by later addition of CC-like materials (*32*). Another possibility is that volatile elements could have been delivered predominantly by NCs, which are thought to be the major building blocks of Earth (*33*). Our results show that Earth has an ε^{40}K value overlapping with those of NCs and CMs; however, CMs cannot be the dominant source of Earth's K because their K concentrations would require Earth to have an implausibly high fraction of CM material (~50% of Earth's mass), which contradicts constraints from several other isotopic systems (*33*).

We calculated the allowed fraction of Earth's K that could have been derived from CCs by assuming that it is from a mixture of NC and CC materials [Fig. 3 (*14*)]. We found that CCs contribute <20% of Earth's K, equivalent to a

maximum contribution of ~13% to Earth by mass over all elements (assuming no volatile loss during or after delivery). This is consistent with previous estimates of several percent to 10% by mass (*33*, *34*), suggesting that Earth's K has a predominant NC origin. The small contribution of CCs for K and other elements (*33*, *34*) is incompatible with a pebble accretion origin of rocky planets (*5*, *35*), but is consistent with the hypothesis that dust-drift barriers separated the CC and NC reservoirs from each other (*25*, *26*, *34*).

REFERENCES AND NOTES

1. B. J. Wood, D. J. Smythe, T. Harrison, *Am. Mineral.* **104**, 844–856 (2019).
2. H.-W. Chen, T. Lee, D.-C. Lee, J. J.-S. Shen, J.-C. Chen, *Astrophys. J. Lett.* **743**, L23 (2011).
3. F. Niederer, D. Papanastassiou, G. Wasserburg, *Geochim. Cosmochim. Acta* **49**, 835–851 (1985).
4. A. Trinquier, J.-L. Birck, C. J. Allegre, *Astrophys. J.* **655**, 1179–1185 (2007).
5. M. Schiller, M. Bizzarro, J. Siebert, *Sci. Adv.* **6**, eaay7604 (2020).
6. R. C. Steele, T. Elliott, C. D. Coath, M. Regelous, *Geochim. Cosmochim. Acta* **75**, 7906–7925 (2011).
7. N. Dauphas, B. Marty, L. Reisberg, *Astrophys. J.* **565**, 640–644 (2002).
8. J. H. Chen, D. A. Papanastassiou, G. J. Wasserburg, *Geochim. Cosmochim. Acta* **74**, 3851–3862 (2010).
9. M. Fischer-Gödde, C. Burkhardt, T. S. Kruijer, T. Kleine, *Geochim. Cosmochim. Acta* **168**, 151–171 (2015).
10. F. Moynier, N. Dauphas, F. A. Podosek, *Astrophys. J.* **700**, L92–L95 (2009).
11. Y. Ku, S. B. Jacobsen, *Sci. Adv.* **6**, eabd0511 (2020).
12. N. X. Nie *et al.*, *Sci. Adv.* **7**, eabl3929 (2021).
13. D. S. Burnett, H. J. Lippolt, G. J. Wasserburg, *J. Geophys. Res.* **71**, 1249–1269 (1966).
14. Materials and methods are available as supplementary materials.
15. G. Srinivasan, A. A. Ulyanov, J. N. Goswami, *Astrophys. J.* **431**, L67–L70 (1994).
16. D. D. Clayton, *Earth Planet. Sci. Lett.* **36**, 381–390 (1977).
17. D. Clayton, *Handbook of Isotopes in the Cosmos: Hydrogen to Gallium* (Cambridge Univ. Press, 2007).
18. T. Rauscher, A. Heger, R. D. Hoffman, S. E. Woosley, *Astrophys. J.* **576**, 323–348 (2002).
19. G. Wasserburg, M. Busso, R. Gallino, K. Nollett, *Nucl. Phys. A.* **777**, 5–69 (2006).
20. G. Jörg, Y. Amelin, K. Kossert, C. L. v. Gostomski, *Geochim. Cosmochim. Acta* **88**, 51–65 (2012).
21. S. Amari, E. Zinner, R. S. Lewis, *Astrophys. J.* **470**, L101–L104 (1996).
22. K. Kossert, E. Günther, *Appl. Radiat. Isot.* **60**, 459–464 (2004).
23. P. H. Warren, *Earth Planet. Sci. Lett.* **311**, 93–100 (2011).
24. T. Kleine *et al.*, *Space Sci. Rev.* **216**, 55 (2020).
25. T. S. Kruijer, C. Burkhardt, G. Budde, T. Kleine, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6712–6716 (2017).
26. T. Lichtenberg, J. Dra Żkowska, M. Schönbächler, G. J. Golabek, T. O. Hands, *Science* **371**, 365–370 (2021).
27. N. Dauphas, A. M. Davis, B. Marty, L. Reisberg, *Earth Planet. Sci. Lett.* **226**, 465–475 (2004).
28. G. K. Nicolussi *et al.*, *Geochim. Cosmochim. Acta* **62**, 1093–1104 (1998).
29. C. Burkhardt *et al.*, *Earth Planet. Sci. Lett.* **312**, 390–400 (2011).
30. C. Burkhardt, T. Kleine, N. Dauphas, R. Wieler, *Earth Planet. Sci. Lett.* **357–358**, 298–307 (2012).
31. C. M. O'D. Alexander, *Geochim. Cosmochim. Acta* **254**, 277–309 (2019).
32. F. Albarède, *Nature* **461**, 1227–1233 (2009).
33. N. Dauphas, *Nature* **541**, 521–524 (2017).
34. C. Burkhardt *et al.*, *Sci. Adv.* **7**, eabj7601 (2021).
35. A. Johansen *et al.*, *Sci. Adv.* **7**, eabc0444 (2021).
36. K and Mo nucleosynthesis calculations for: N. X. Nie, D. Wang, Z. A. Torrano, R. W. Carlson, C. M. O'D. Alexander, A. Shahar, Meteorites have inherited nucleosynthetic anomalies of potassium-40 produced in supernovae, Dryad (2022); https://doi.org/10.5061/dryad.brv15dvcz.

ACKNOWLEDGMENTS

We thank R. Trappitsch, N. Liu, D.S. Grewal, F.L.H. Tissot, P. Ni, L.R. Nittler, T. Hopp, Z. Zhang, and N. Dauphas for helpful discussions; P. Heck of the Robert A. Pritzker Center for Meteoritics and Polar Studies of the Field Museum, the Muséum National d'Histoire Naturelle, and the Smithsonian Natural History Museum for providing the meteorite samples; and T. Mock for assistance with the TIMS instrument. **Funding:** N.X.N. was supported by a NASA NESSF fellowship (NNX15AQ97H), a Carnegie postdoctoral fellowship, and a Carnegie Postdoc × Postdoc (P^2) seed grant. D.W. was supported by a Carnegie postdoctoral fellowship. **Author contributions:** N.X.N., D.W., R.W.C., and A.S. initiated the study. R.W.C. and C.M.O'D.A. provided samples. N.X.N. and D.W. performed sample digestion and chemical purification of K. D.W. and N.X.N. developed the TIMS measurement method and measured the K isotopic compositions of the samples with help from R.W.C. and A.S. N.X.N., D.W., and Z.A.T. measured the elemental abundances of the samples. All authors contributed to discussions and interpretations of the data. N.X.N. wrote the initial draft of the manuscript, which was subsequently edited by D.W., Z.A.T., R.W.C., C.M.O'D.A., and A.S. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All raw isotope measurements are provided in data S1, and the derived ε^{40}K values are listed in table S1. The sources of all meteorites and terrestrial standards are specified in the supplementary materials. The K and Mo nucleosynthesis calculations are available at Dryad (*36*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn1783
Materials and Methods
Supplementary Text
Figs. S1 to S10
Table S1
References (*37*–*126*)
Data S1

Submitted 10 November 2021; accepted 16 December 2022
10.1126/science.abn1783

NEUROSCIENCE

Encoding of environmental illumination by primate melanopsin neurons

Andreas Liu[1], Elliott S. Milner[1]†**, Yi-Rong Peng[2]‡, Hannah A. Blume[1], Michael C. Brown[1], Gregory S. Bryman[1]§, Alan J. Emanuel[1]#, Philippe Morquette[1], Nguyen-Minh Viet[1], Joshua R. Sanes[2], Paul D. Gamlin[3], Michael Tri H. Do[1]***

Light regulates physiology, mood, and behavior through signals sent to the brain by intrinsically photosensitive retinal ganglion cells (ipRGCs). How primate ipRGCs sense light is unclear, as they are rare and challenging to target for electrophysiological recording. We developed a method of acute identification within the live, ex vivo retina. Using it, we found that ipRGCs of the macaque monkey are highly specialized to encode irradiance (the overall intensity of illumination) by blurring spatial, temporal, and chromatic features of the visual scene. We describe mechanisms at the molecular, cellular, and population scales that support irradiance encoding across orders-of-magnitude changes in light intensity. These mechanisms are conserved quantitatively across the ~70 million years of evolution that separate macaques from mice.

Many visual functions are non–image-forming: they respond less to details in the scene than to irradiance. For example, irradiance indicates the Sun's position in the sky and is therefore used to synchronize the circadian clock with local time, regulating patterns of physiology and gene expression in practically every tissue (*1*). IpRGCs are necessary for circadian photoregulation and many other non-image visual functions (*2*). These neurons sense light directly using a G protein–coupled receptor called melanopsin and indirectly through retinal circuitry (*2*, *3*). We investigated how primate ipRGCs encode irradiance, focusing on two key requirements: insensitivity to image detail and the signaling of light levels spanning several orders of magnitude from night to day (*1*, *4*). Our study rests on two technical advances: First, we developed an immunotagging-based approach that identifies these rare neurons acutely in the live, ex vivo retina, making them accessible to single cell analysis. Second, we transplanted macaque melanopsin into mouse ipRGCs, allowing properties of this molecule to be examined in a well-defined context. Our experiments concern the intrinsic, melanopsin-driven light response that distinguishes ipRGCs from other retinal output neurons (*2*, *5*).

[1]F. M. Kirby Neurobiology Center and Department of Neurology, Boston Children's Hospital and Harvard Medical School. Boston, MA 02115, USA. [2]Department of Molecular and Cellular Biology and Center for Brain Science, Harvard University, Cambridge, MA 02138, USA. [3]Department of Ophthalmology and Visual Sciences, University of Alabama at Birmingham, Birmingham, AL, 35294, USA.
***Corresponding author. Email: michael.do@childrens.harvard.edu**
†Present address: Sainsbury Wellcome Centre for Neural Circuits and Behavior, University College London, London, W1T 4JG, UK.
‡Present address: Department of Ophthalmology, Stein Eye Institute, UCLA David Geffen School of Medicine, Los Angeles, CA 90095, USA.
§Present address: Merck & Co., Inc., Cambridge, MA 02141, USA.
#Present address: Department of Cell Biology, Emory University School of Medicine, Atlanta, GA, 30322, USA.
**These authors contributed equally to this work.

Ex vivo, live identification of macaque ipRGCs

In humans and other primates, ipRGCs constitute ~0.4% of the retinal output neurons (*6*) and require a reliable method of identification for systematic study. To examine their electrical responses, live identification is needed (*3*, *7*).

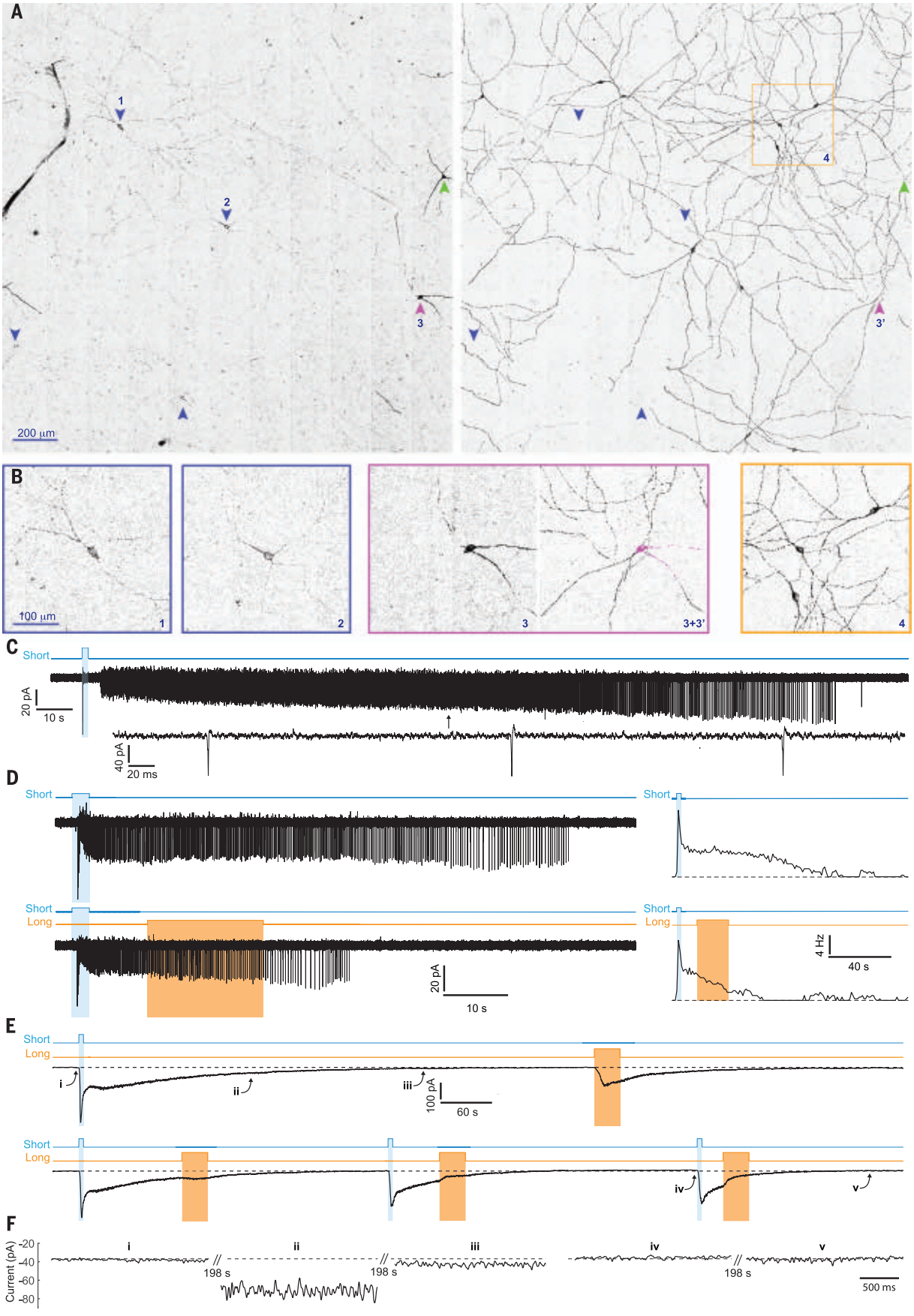

A
1
2
3
4
3'
200 µm
B
100 µm
1
2
3
3+3'
4
C
Short
20 pA
10 s
40 pA
20 ms
D
Short
Long
20 pA
10 s
4 Hz
40 s
E
Short
Long
i
ii
iii
100 pA
60 s
iv
v
F
Current (pA)
-20
-40
-60
-80
198 s
198 s
198 s
500 ms

Figure 1

Fig. 1. Photoswitchable persistent responses in live-immunotagged macaque ipRGCs. (**A**) Antibody labeling in live tissue reveals diverse ipRGCs (inverted fluorescence images). (Left) Z projection through a 13.5-μm thick volume containing the ganglion cell layer (GCL) and inner aspect of the inner plexiform layer (IPL). Arrowheads mark ipRGCs with somata in the GCL. Fluorescence in the upper left is from a blood vessel. (Right) for the same XY position, a Z projection through a 29-μm thick volume containing the rest of the IPL and the inner aspect of the inner nuclear layer (INL). Arrowheads are in the same positions. (**B**) (Left to right) expanded views of two lightly stained ipRGCs with somata in the GCL; an outer-stratifying ipRGC with its soma in the GCL (superficial and deep projections juxtaposed); and three outer-stratifying ipRGCs with somata displaced to the INL [corresponding to the orange box in (A)]. (**C**) Loose patch recording. A pulse of short-wavelength light evokes persistent spiking near body temperature (35°C). The inset shows spike waveforms on an expanded time base (excerpt position marked with an arrow) (**D**) (Top) Persistent spiking in another ipRGC. Peristimulus time histogram (PSTH) for three trials shown on right (23°C). (Bottom) For the same cell, chasing the short-wavelength light with long-wavelength light curtails the persistent response (PSTH from three trials). (**E**) Perforated-patch recording in voltage clamp (−80 mV, 23°C). Short- and long-wavelength switching of persistent photocurrent. Long-wavelength light evokes current because some photon absorptions cycle melanopsin molecules through the signaling state, even as the photoequilibrium favors low occupancy of this state. Response traces are low-pass filtered at 2 Hz. (**F**) 2-s long segments of the same recording, keyed to time points in (E). Traces are low-pass filtered at 20 Hz. Synaptic antagonists. (For all figures, see Methods for stimulus details.)

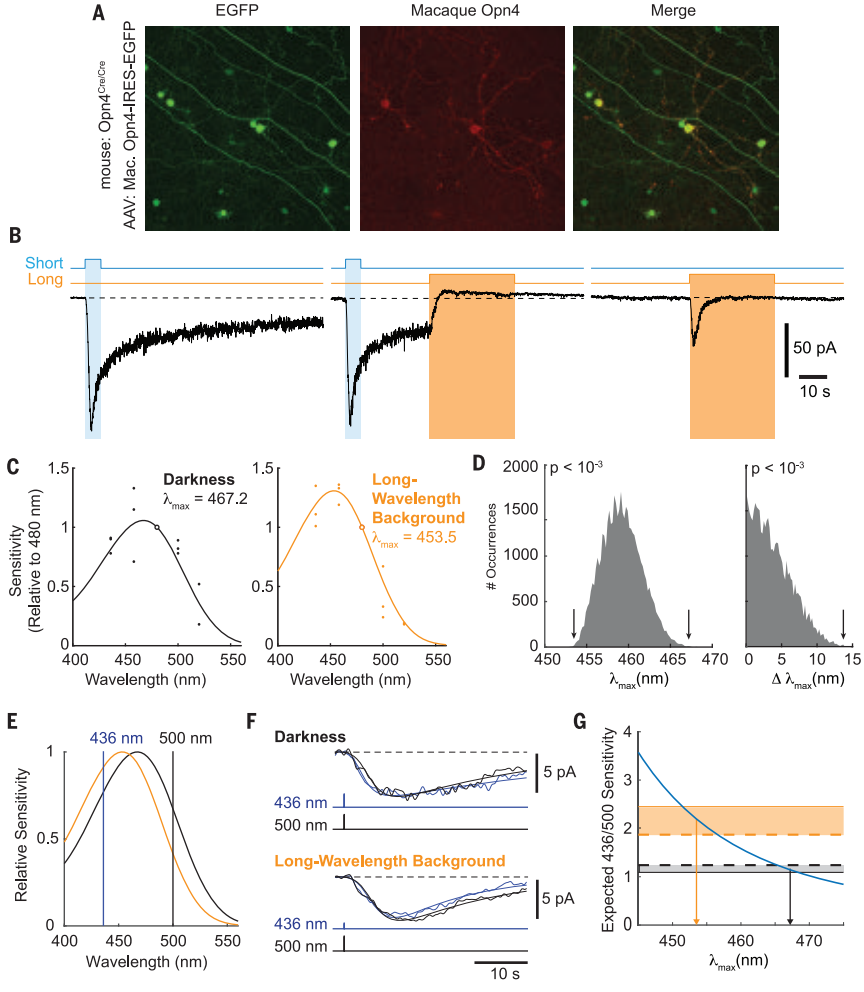


Fig. 2. Macaque melanopsin multistability underlies temporal and chromatic integration. (**A**) "Hybrid" mouse ipRGCs have both melanopsin alleles replaced by Cre recombinase (*Opn4*$^{Cre/Cre}$) and are transduced with a virus driving Cre-dependent macaque melanopsin and EGFP. Immunolabeling for EGFP and macaque melanopsin are shown with a merge. (**B**) Short- and long-wavelength switching of persistent phototransduction in a hybrid ipRGC. (**C**) Criterion action spectra measured in hybrid ipRGCs (Methods), reflecting the absorption spectra of macaque melanopsin, following prolonged darkness (left; 11 cells) or from atop a background of 600-nm light (right; 11 different cells). The data are fit to templates (*18*) for single pigment states, yielding peak wavelength sensitivity (λ_{max}) values of 467.2 nm (darkness) and 453.5 nm (600-nm background). Each curve is fit to two cells tested at 520 nm and three cells tested at each of three wavelengths: 436, 458, and 500 nm. Sensitivities at these wavelengths are normalized to the sensitivity at 480 nm (open circles). (**D**) (Left) Permutation test giving probabilities of observing λ_{max} >467.2 nm or <453.5 of 3.96×10^{-4} and 3.54×10^{-4}, respectively. (Right) Permutation test showing a small probability ($P = 5.00 \times 10^{-4}$) of observing a spectral separation between the two states of >13.7 nm by chance. Arrows indicate actual best-fit values. (**E**) Expected sensitivity to 436- and 500-nm light in darkness (ratio of 1.14) or atop a 600-nm background (ratio of 2.20) given the melanopsin states defined in (C). (**F**) Criterion responses evoked from a macaque ipRGC by 436- and 500-nm flashes under these adaptation conditions. Responses are fit using an equation with a sigmoidal rise and an exponential decay (Methods). Heights of flash monitors indicate relative stimulus intensity (photon density) at the two wavelengths. (**G**) A pigment state with a given λ_{max} (x axis) has an expected 436/500-nm sensitivity ratio (y axis) that is denoted by the blue curve. The λ_{max} values exhibited by hybrid ipRGCs under dark adaptation and 600-nm adaptation are indicated by black and orange vertical arrows, respectively. The 436/500-nm sensitivity ratios measured for two macaque ipRGCs under dark or 600-nm adaptation are shown as black and orange horizontal lines, respectively [the solid line represents the cell in (F) and the dashed line another cell; measurement ranges are shaded]. Synaptic antagonists, 23°C.

We sought an ex vivo method that could be applied acutely to live tissue collected opportunistically. We raised an antibody against melanopsin's N terminus. The antibody labeled live cells with the morphological features of ipRGCs (*3*, *6*) (Fig. 1, A and B). We further validated the antibody using heterologous expression of macaque melanopsin in cell lines and mouse ipRGCs (Fig. 2 and figs. S1 and S2).

We compared light responses between macaque ipRGCs (identified with antibody) and mouse ipRGCs expressing macaque melanopsin in place of their own (identified without antibody) and detected no differences in melanopsin function (see below). This method of identifying cells within live tissue should generalize to other cell types of any species for which ectodomain markers are accessible.

Our study concerns outer-stratifying ipRGCs, whose dendrites are placed deep within the retina (henceforth, simply "ipRGCs"; see Methods). They are anatomical orthologs of the M1 type of rodent ipRGCs. M1 ipRGCs have the highest melanopsin expression, exhibit the largest intrinsic light responses, and are essential drivers of non-image visual functions that include circadian photoregulation (*2*, *8*, *9*).

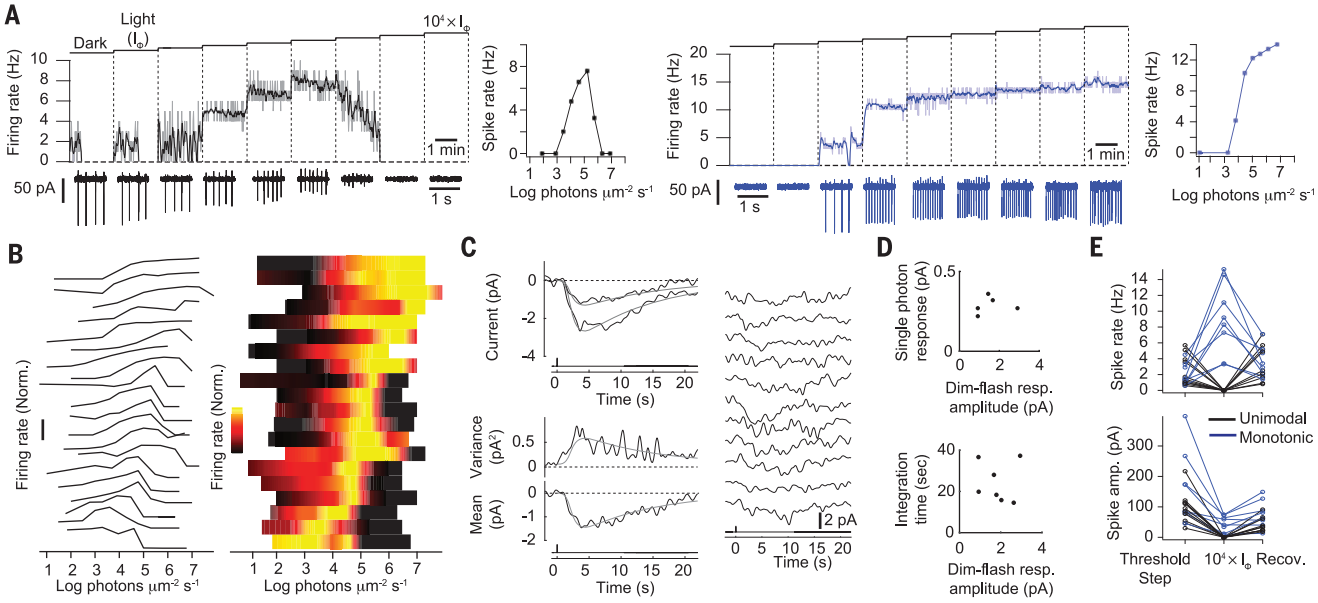


Fig. 3. Macaque ipRGCs collectively encode a broad range of light intensities through differential intrinsic tuning. (**A**) Loose-patch recordings from macaque ipRGCs. (Left) As light intensity increased, this cell increased spiking and then silenced. Its tuning curve (intensity-firing relation) is unimodal (right). Spikes diminished in size until silencing. Spiking resumed when light intensity decreased (fig. S6A). (Right) This cell increased spiking toward a plateau and did not silence. Its tuning curve is monotonic. (**B**) Unimodal and monotonic ipRGCs are tuned to different irradiances. Curves are normalized by their peaks and vertically offset for display. (**C**) Single photon response estimated using perforated-patch, voltage-clamp recordings (−80 mV) from macaque ipRGCs. (Upper left) Linear superposition of dim-flash responses in an ipRGC. When scaled by stimulus intensity, the same smooth curve (see Methods) fits both responses. Traces are averages of 16 trials, low-pass filtered at 0.5 Hz. (Lower left) Variance and mean of the same ipRGC's responses to 11 identical dim flashes (trials shown at right). The variance and the mean can be fit by the same smooth curve, scaled differently. Dividing the variance by the mean gives the estimated single photon response amplitude, which is 0.36 pA for this example. (Right) Individual trials for this cell. (**D**) The single photon response amplitude (variance/mean ratio) and the integration time (integral/peak ratio) computed in 5 and 7 ipRGCs, respectively. These parameters did not depend on the amplitudes of the responses analyzed, whose means ranged from 0.92 to 2.94 pA. (**E**) Many ipRGCs are tuned by depolarization block. (Top) Spike rates at the threshold (I_Φ; Methods) and top ($10^4 \times I_\Phi$) irradiances. "Recov." is the spike rate closest to the threshold rate recovered when irradiance is lowered. (Bottom) Same analysis but for recovery of spike amplitude. Synaptic antagonists, 35°C (A, B, and E) or 23°C (C and D).

To focus on cell-autonomous mechanisms, we applied antagonists of synaptic transmission in all electrophysiological experiments.

Macaque melanopsin supports temporal and spatial integration

To study how macaque ipRGCs encode irradiance, we examined the electrical spikes that they fire to convey visual information to the brain. Even when driven by their intrinsic responses alone, ipRGCs spiked during illumination and long afterward; post-illumination, persistent firing lasted 270 ± 50 s (mean ± SD, 3 cells examined at 35°C; Fig. 1C). Macaques make large eye movements several times per second to scan a given visual scene (*10*). Persistent spiking continues over more than a thousand such saccades, blurring spatial information and favoring irradiance encoding.

We asked how persistent spiking is driven by melanopsin, a pigment of the rhabdomeric type (*2*). These pigments tend to be bistable: when the ground state absorbs a photon, it switches to an active state that has a long lifetime (*11*). When the active state absorbs a photon, it switches to the ground state, deactivating the cell. Evidence for bistability has been reported for the melanopsin of primates (*12*) and the invertebrate amphioxus (*13*). These studies suggest that long-wavelength light (≥ ~560 nm) is preferentially absorbed by melanopsin's active state and should therefore produce deactivation, although this has not been tested directly in primate ipRGCs. We found that long-wavelength illumination acutely suppressed the persistent spiking of macaque ipRGCs (Fig. 1D and fig. S3; 3 of 3 cells). Thus, persistent spiking appears to originate with the long-lived active state of a photoswitchable melanopsin.

Because persistent activity can be mediated by voltage-gated ion channels (*14*), we clamped the membrane voltage to suppress these channels. Photoswitchable persistent responses remained (Fig. 1, E and F; 3 of 3 cells, decay time constant 123 to 404 s at 23°C), providing additional evidence that they arise from melanopsin.

Unlike the monostable pigments of rods and cones, which disassemble after activation and require accessory cells for reassembly (*2*), a melanopsin molecule that has been switched off by light is available immediately to be switched on again. This supports a sustained representation of irradiance. Indeed, without melanopsin some non-image visual functions are abnormally transient (*2*).

Macaque melanopsin supports a degree of wavelength integration

Our experiments indicate that melanopsin can be switched from an active to an electrically silent state but have not addressed the nature of that silent state. If melanopsin is bistable, that state is the ground state. However, mouse melanopsin appears tristable, possessing a ground state, an active state, and an "extra" state (E) that is electrically silent (*15*, *16*). The ground and active states absorb long wavelengths more effectively than the E state, so that long-wavelength light causes the E state to accumulate. We asked whether macaque melanopsin deactivates by returning to the ground state (consistent with bistability) or by switching to another electrically silent state.

We examined macaque melanopsin in mouse ipRGCs that lacked their endogenous melanopsin alleles through replacement with Cre recombinase (*17*). These "hybrid" ipRGCs (mouse ipRGCs expressing macaque melanopsin) are relatively plentiful and provide a more physiological setting than immortalized cell lines

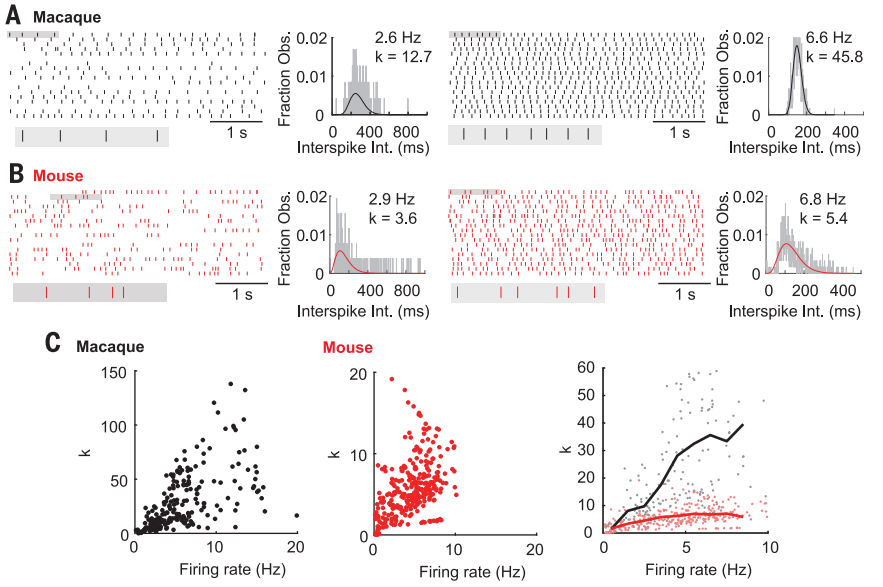


Fig. 4. Regular spiking of macaque ipRGCs. (**A**) Spiking of a macaque ipRGC at low and intermediate firing rates. Insets display shaded portions of spike trains on an expanded time base. Spike times and histograms of interspike intervals (ISIs) are shown. Fits are gamma distributions yielding the regularity parameters (k) indicated. (**B**) The same for a mouse ipRGC. (**C**) Values of k for ISIs in 20 macaque ipRGCs (black) and 26 mouse ipRGCs (red). Each cell is represented by 5 to 16 data points corresponding to stimulus epochs with different firing rates. The correlation of k with firing rate across stimulus epochs for an individual cell was 0.70 ± 0.15 (macaque) or 0.67 ± 0.21 (mouse). Solid lines are averages over 1-Hz bins. For a given firing rate, k has a higher average value in macaque cells than in mouse cells. Saturated (95% maximal) firing rates are similar in both species (see main text). Synaptic antagonists, 35°C.

and biochemical preparations (Fig. 2A and fig. S2). We recorded hybrid ipRGCs that are likely of the M1 type (*2*) (see Methods). Macaque melanopsin activates phototransduction in these neurons. Similar to macaque ipRGCs, hybrid ipRGCs produced persistent responses that were suppressed by long-wavelength light (Fig. 2B and fig. S4, A to C; 12 of 12 cells). We therefore used hybrid cells to study the electrically silent state induced by long-wavelength illumination.

We measured the wavelength sensitivity of cellular photoactivation (*11*, *15*, *18*). If macaque melanopsin were bistable, the action spectrum should be constant because it always reflects the absorption spectrum of the single ground state. However, this was not the case (Fig. 2C and fig. S4D). Following prolonged darkness to produce the ground state, we measured an action spectrum whose peak wavelength sensitivity (λ_{max}; see Methods) was 467.2 nm. During long-wavelength illumination, the action spectrum had a λ_{max} of 453.5 nm instead, matching the E state of mouse melanopsin (*15*, *16*) and indicating the presence of a second electrically silent state of macaque melanopsin. Statistical analyses indicate that the probability of our observing such widely separated values of λ_{max} by chance is small ($P < 10^{-3}$; Fig. 2D). We also obtained quantitatively similar results in macaque ipRGCs (Fig. 2, E to G, and fig. S5). Thus, our experiments indicate that macaque melanopsin activates from at least two electrically silent states whose preferred wavelengths appear cyan (467.2 nm) and violet (453.5 nm) to trichromatic humans. This spectral separation is roughly half of that between the two cone pigments that define the red-green axis of color vision in humans and macaques (*19*). Under natural light, melanopsin molecules could activate from both silent states (*15*), limiting the cell's wavelength selectivity to promote irradiance encoding.

Because we identified hybrid ipRGCs without immunotagging, our experiments with these cells provide a natural control for effects of the procedure and indicate that it is innocuous for melanopsin function. Following live immunotagging, macaque melanopsin's signaling state retains its thermal stability and its ability to be switched off by long-wavelength light (compare Fig. 1, C to F, with Fig. 2B and fig. S4, A to C), and the silent states retain their distinct spectral sensitivities (compare Fig. 2, C and E, with Fig. 2, F and G and fig. S5B).

Population coding of light intensity by macaque ipRGCs

The intensity of environmental illumination varies over many orders of magnitude (*4*), whereas individual neurons have a much smaller range of distinguishable spike rates. We asked how the intrinsic mechanisms of primate ipRGCs help meet this challenge. We first determined the range over which melanopsin phototransduction drives spikes in these neurons. Rising irradiance caused spiking to increase and then silence for many macaque ipRGCs (11 of 20 cells; Fig. 3, A and B, and fig. S6), giving unimodal tuning curves (intensity-firing relations). For the remainder, spiking increases to a plateau or decreases slightly, giving monotonic tuning curves. Individually, ipRGCs saturated at relatively low firing rates (95% maximal spike rate of 8.2 ± 3.3 Hz; 20 cells) and modulated their spiking over a range of 2.7 ± 0.83 log units of light intensity (19 of 20 cells; 1 excluded because threshold and peak firing rates were similar). Collectively, ipRGCs spanned a range of 5.2 log units, >100 times broader. The most sensitive cells were activated at 1.4 log photons μm^{-2} sec^{-1}, approximately the intensity of moonlight and comparable to the threshold of cone photoreceptors (*20*). The least sensitive were saturated at 6.7 log photons μm^{-2} sec^{-1}, approximately the intensity of full daylight. Comparing half-saturating irradiances, the range is 2.5 log units. Rod and cone inputs are likely to extend this range even further (*3*).

We next asked how some ipRGCs exhibit cone-like sensitivity. High photon capture by ipRGCs would interfere with image vision because ipRGCs lie ahead of rods and cones in the eye's light path. We hypothesized that, as in nocturnal mice (*21*), ipRGC sensitivity in diurnal primates is increased by the properties of the response to each captured photon. This single photon response originates with one photoactivated melanopsin molecule and is a building block of phototransduction. To define it, we examined the responses to dim flashes. We found linear superposition of small responses (generally <3 pA at room temperature, used for additional recording stability; Fig. 3C, fig. S7; Methods). The variance and mean of linear responses can be used to estimate the amplitude and duration (integration time) of the single photon response (*21*). These parameters were 0.29 ± 0.04 pA and 24 ± 8 s, respectively, when measured at room temperature (Fig. 3D). Extrapolating to body temperature (*21*, *22*) (see Methods), where we measured tuning curves, yields ~0.87 pA and ~11 s. These values exceed those of cone single photon responses, which have an amplitude of ~40 fA and an integration time of ~70 ms (*23*). The large size and long lifetime of the ipRGC single photon response favor sensitivity, the latter by increasing the likelihood that responses to successive photon captures will sum.

Behavioral studies indicate that melanopsin phototransduction is especially critical in bright light (*2*). The subset of monotonically tuned ipRGCs is important in this regime; each unimodally tuned ipRGC is silenced as the light intensity exceeds its preferred range. Unimodal

ipRGCs tend to be those that are most sensitive and thus reach their maximum firing rates under relatively low light (intensities at peak firing for unimodal and monotonic ipRGCs are 4.8 ± 0.74 and 6.0 ± 0.62 log photons μm^{-2} s^{-1}, respectively; $P < 10^{-3}$). Without silencing, these cells would fire at their maximum rates through much of the day. Silencing these relatively uninformative ipRGCs is likely to promote energetic efficiency (*24*). Monotonic ipRGCs are efficient because they fire at their highest rates only in the brightest light, which lasts a comparatively short period of time as the Sun follows its arc.

How does silencing occur? As firing rates fall with increasing irradiance, spikes dwindle in height. This is a signature of depolarization block, where the spike generator is overdriven into a refractory state (Fig. 3, A and E). Depolarization block is increasingly recognized as a physiological state (*25*–*27*), and our data reveal that it tunes primate neurons.

Discussion

In this study, we developed a live immunotagging technique to study a defined subpopulation of macaque neurons. This technique should generalize to other cell types and species. We used it to locate outer-stratifying macaque ipRGCs for functional analysis. Concurrently, we used heterologous expression to study properties of macaque melanopsin in the background of mouse ipRGCs. These approaches revealed specializations at the molecular, cellular, and population scales that allow ipRGCs to encode environmental illumination.

Our results on the outer-stratifying macaque ipRGCs permit comparisons with the orthologous population of outer-stratifying mouse ipRGCs (i.e., M1 ipRGCs) (*15*, *21*, *22*, *26*, *28*). A notable divergence lies in spiking regularity. Interspike intervals (ISIs) in macaque ipRGCs have a coefficient of variation (CV) whose minimum is 0.23 ± 0.09 (20 cells). Fitting ISI distributions with gamma functions (*29*) yields a maximum regularity parameter (k) of 60.8 ± 29.4, more than 3 times higher than the maximum value observed in a survey of macaque cortical neurons (*29*) (Fig. 4, A and C, and fig. S8). Analyzing existing data from mouse ipRGCs (*26*), we find minimum CV and maximum k values of 0.71 ± 0.19 and 8.9 ± 2.6 (26 cells), indicating 3.1 and 6.9 times lower spike regularity than that of macaque ipRGCs, respectively (Fig. 4, B and C, and fig. S8). These differences are not a consequence of different spike rates between macaque and mouse ipRGCs (95% maximal rates of 8.2 ± 3.3 and 7.7 ± 0.9 Hz, respectively; Fig. 4C). In principle, the regular spiking of macaque ipRGCs allows different irradiance levels to be distinguished more readily (fig. S9).

With the exception of regular firing, our study indicates that cell-autonomous mechanisms are largely conserved between mouse and macaque outer-stratifying ipRGCs. Similar to macaque melanopsin, mouse melanopsin has a long-lived signaling state that supports temporal (and hence spatial) integration as well as spectrally distinct silent states that support a degree of chromatic integration (*15*). Similar to macaque ipRGCs, mouse ipRGCs produce a large and prolonged single photon response (*21*, *22*). Mouse ipRGCs also exhibit differential tuning to irradiance (*22*, *26*), resembling macaque ipRGCs in the dynamic ranges of both single cells and the population.

Our findings have the potential to improve artificial lighting, including that used for therapy in mood disorders (*5*). For example, we observe that moderate and not bright light causes the largest number of ipRGCs to signal downstream and that certain wavelengths suppress the persistent activity of these neurons. Future studies should examine how these intrinsic properties combine with rod- and cone-driven signals to shape the outputs of ipRGCs (*2*, *3*, *5*).

Organisms as diverse as coral and humans sense light using melanopsin (*30*). The melanopsin-expressing ipRGCs exist in every mammal so far examined, even subterranean mole rats (*2*). Mice and macaques are separated by ~70 million years of evolution (*31*) as well as by their divergent visual ecologies (e.g., nocturnal versus diurnal). Nonetheless, the functional requirements of irradiance encoding—insensitivity to image detail and the need for a wide dynamic range—are constant across the divide. The persistence of distinctive molecular, cellular, and population-level features suggests that these functional requirements place strong constraints on neurons that have evolved to report the intensity of environmental illumination.

REFERENCES AND NOTES

1. J. W. Mouland, A. R. Stinchcombe, D. B. Forger, T. M. Brown, R. J. Lucas, *Curr. Biol.* **27**, 1633–1640.e3 (2017).
2. M. T. H. Do, *Neuron* **104**, 205–226 (2019).
3. D. M. Dacey *et al*., *Nature* **433**, 749–754 (2005).
4. F. Rieke, M. E. Rudd, *Neuron* **64**, 605–616 (2009).
5. R. J. Lucas *et al*., *Trends Neurosci.* **37**, 1–9 (2014).
6. H.-W. Liao *et al*., *J. Comp. Neurol.* **524**, 2845–2872 (2016).
7. P. D. Gamlin *et al*., *Vision Res.* **47**, 946–954 (2007).
8. A. D. Güler *et al*., *Nature* **453**, 102–105 (2008).
9. M. Hatori *et al*., *PLOS ONE* **3**, e2451 (2008).
10. A. L. Yarbus, *Eye Movements and Vision* (Plenum Press, 1967).
11. P. Hillman, S. Hochstein, B. Minke, *Physiol. Rev.* **63**, 668–772 (1983).
12. L. S. Mure *et al*., *PLOS ONE* **4**, e5991 (2009).
13. M. Koyanagi, K. Kubokawa, H. Tsukamoto, Y. Shichida, A. Terakita, *Curr. Biol.* **15**, 1065–1069 (2005).
14. J. Zylberberg, B. W. Strowbridge, *Annu. Rev. Neurosci.* **40**, 603–627 (2017).
15. A. J. Emanuel, M. T. H. Do, *Neuron* **85**, 1043–1055 (2015).
16. T. Matsuyama, T. Yamashita, Y. Imamoto, Y. Shichida, *Biochemistry* **51**, 5454–5462 (2012).
17. J. L. Ecker *et al*., *Neuron* **67**, 49–60 (2010).
18. V. I. Govardovskii, N. Fyhrquist, T. Reuter, D. G. Kuzmin, K. Donner, *Vis. Neurosci.* **17**, 509–528 (2000).
19. S. L. Merbs, J. Nathans, *Nature* **356**, 433–435 (1992).
20. N. T. Ingram, A. P. Sampath, G. L. Fain, *J. Physiol.* **594**, 5415–5426 (2016).
21. M. T. H. Do *et al*., *Nature* **457**, 281–287 (2009).
22. A. J. Emanuel, K. Kapur, M. T. H. Do, *Cell Rep.* **21**, 1048–1062 (2017).
23. L. H. Cao, D. G. Luo, K. W. Yau, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2752–2757 (2014).
24. P. Sterling, S. Laughlin, *Principles of Neural Design* (MIT Press, 2015).
25. M. D. Belle, C. O. Diekman, D. B. Forger, H. D. Piggins, *Science* **326**, 281–284 (2009).
26. E. S. Milner, M. T. H. Do, *Cell* **171**, 865–876.e16 (2017).
27. S. Wienbar, G. W. Schwartz, *Neuron* **110**, 2110–2123.e4 (2022).
28. M. T. H. Do, K.-W. Yau, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 7470–7475 (2013).
29. G. Maimon, J. A. Assad, *Neuron* **62**, 426–440 (2009).
30. M. T. H. Do, K.-W. Yau, *Physiol. Rev.* **90**, 1547–1581 (2010).
31. T. R. Disotell, A. J. Tosi, *Genome Biol.* **8**, 226 (2007).

ACKNOWLEDGMENTS

We thank those who shared materials: J. Assad, R. Born, C. Cetrulo, J. DiCarlo, A. Hall, R. P. Johnson, M. Livingstone, J. Madsen, T. Moore, M. Nedelman, M. Papisov, D. Rosene, and S. Smith. We thank R. Born and M. Greenberg for helping to initiate and sustain this research program; V. Berezovskii, E. Cosgrove, E. Curran, E. Egan, and R. Adams for logistical and technical assistance; C. Wang, Y. Zhang, and Z. He for viral vectors; F. Rieke and R. Sinha for expertise on the primate retina; C. Cepko and W. Xiong for expertise on molecular biology; D. Vavvas for microsurgical techniques; S. Hattar for the $Opn4^{Cre}$ mouse line; D. Felis, W. Lorenzen, and J. Wu for institutional infrastructure; and F. Caval-Holme for comments on the manuscript. M. T. H. Do is affiliated with the Center for Brain Science of Harvard University, the Division of Sleep Medicine of Harvard Medical School, and the Broad Institute of MIT and Harvard. **Funding:** This work was supported by the following: National Institutes of Health R01EY023648, R01EY030628, R01EY034089, and R21EY025840 (to M.T.H.D.); R01EY025555 (to P.D.G. and M.T.H.D.); R21EY028633 (to J.R.S. and M.T.H.D.); F31EY025466 (to E.S.M.); P50HD105351 (to Boston Children's Hospital IDDRC); P30EY012196 (to Harvard Medical School); The National Science Foundation (GFRP to G.S.B. and A.J.E.); The BrightFocus Foundation M2014055 (to M.T.H.D.); the Lefler Center (to A.L. and P.M.); the Neurobiology Department of Harvard Medical School (to A.L.); the Tommy Fuss Center (to P.M.). **Author contributions:** A.L., G.S.B., A.J.E., and M.T.H.D. developed the method of live immunotagging in retina. A.L., G.S.B., and M.T.H.D. developed methods of obtaining, transporting, and maintaining live tissue. A.L., G.S.B., M.C.B., A.J.E., and M.T.H.D. procured tissue. A.L. obtained perforated-patch electrophysiological data except for those in fig. S4, A to C, obtained by N.-M.V. E.S.M. obtained all loose-patch electrophysiological data except for one cell, contributed by P.M., A.L., E.S.M., and M.T.H.D. analyzed electrophysiological data. Y.-R.P. cloned and subcloned macaque melanopsin in the laboratory of J.R.S.-P.D.G., and M.T.H.D. generated and validated custom antibodies. E.S.M. and H.A.B. performed intraocular injections. M.C.B. and M.T.H.D. conducted immunohistochemical experiments on fixed tissues. A.L. and M.T.H.D. wrote the manuscript with input from all authors. M.T.H.D. conceived of live immunotagging and of exchanging melanopsin species within native neurons. H.A.B., M.C.B., G.S.B., A.J.E., P.M., and N.-M.V. are listed alphabetically. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Materials are available upon request to michael.do@childrens.harvard.edu. Data are stored and curated by Harvard Dataverse (*41*).

License information: Copyright © 2023 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.ade2024
Materials and Methods
Figs. S1 to S9
References (*32*–*41*)
MDAR Reproducibility Checklist

Submitted 2 August 2022; accepted 21 November 2022
10.1126/science.ade2024

GEOCHEMISTRY

How temperature-dependent silicate weathering acts as Earth's geological thermostat

S. L. Brantley[1,2]*, Andrew Shaughnessy[2], Marina I. Lebedeva[1]†, Victor N. Balashov[1]†

Earth's climate may be stabilized over millennia by solubilization of atmospheric carbon dioxide (CO_2) as minerals weather, but the temperature sensitivity of this thermostat is poorly understood. We discovered that the temperature dependence of weathering expressed as an activation energy increases from laboratory to watershed as transport, clay precipitation, disaggregation, and fracturing increasingly couple to dissolution. A simple upscaling to the global system indicates that the temperature dependence decreases to ~22 kilojoules per mole because (i) the lack of runoff limits weathering and retains base metal cations on half the land surface and (ii) other landscapes are regolith-shielded and show little weathering response to temperature. By comparing weathering from laboratory to globe, we reconcile some aspects of kinetic and thermodynamic controls on CO_2 drawdown by natural or enhanced weathering.

Over multimillion-year time scales, the balance between weathering of silicate rocks and volcanic degassing may control the atmospheric concentration of CO_2, one of the most important greenhouse gases that regulate Earth's climate. Silicate weathering accelerates with temperature, acting as a negative feedback that buffers Earth's climate and maintains its habitability (*1*–*5*). Steps by which silicate weathering maintains Earth's low and generally stable atmospheric partial pressure of CO_2 (P_{CO_2}) include (i) production of alkalinity through solubilization of base metal cations and atmospheric CO_2 during silicate weathering, (ii) riverine transport of dissolved inorganic carbon (DIC) to the oceans, and (iii) precipitation and burial of DIC as carbonate minerals at the seafloor. Weathering of marine basalt at increased temperatures also sequesters CO_2 into minerals at the seafloor (*6*). In contrast to silicates, CO_2 drawn from the atmosphere during weathering of carbonate minerals is ultimately rereleased to the atmosphere. Despite the importance of the silicate weathering feedback, the mechanism of its operation is not well understood. Proposed controls vary from chemical equilibrium to kinetics to hydrology, and the sensitivity may be as low as 2% K^{-1} or perhaps as high as 20% K^{-1} [e.g., (*5*, *6*–*13*)]. We explore the temperature sensitivity of weathering to understand how this feedback functions.

Although laboratory measurements might appear to be the gold standard to determine how silicate weathering varies exponentially with temperature, rates are too slow to assess accurately. Furthermore, dissolution is only one of many temperature-dependent steps that occur during weathering. Additional processes—such as solute transport, clay precipitation, biotic activity, disaggregation, fracturing, and erosion—enter systems at larger scales. These processes in turn can be influenced by climate factors such as temperature, runoff, and precipitation and by weatherability factors such as lithology, porosity, permeability, type of vegetation, and landscape position or distribution of land mass [e.g., (*1*, *14*, *15*, *16*)]. Whether kinetics or equilibrium controls the temperature sensitivity remains an open question (*9*).

We hypothesized that combined observations from laboratories, soils, watersheds, and the globe could elucidate puzzles about weathering versus temperature. For example, researchers disagree about how chemical weathering

[1]Earth and Environmental Systems Institute, Pennsylvania State University, University Park, PA, USA. [2]Department of Geosciences, Pennsylvania State University, University Park, PA, USA.
*Corresponding author. Email: brantley@geosc.psu.edu
†These authors contributed equally to this work.

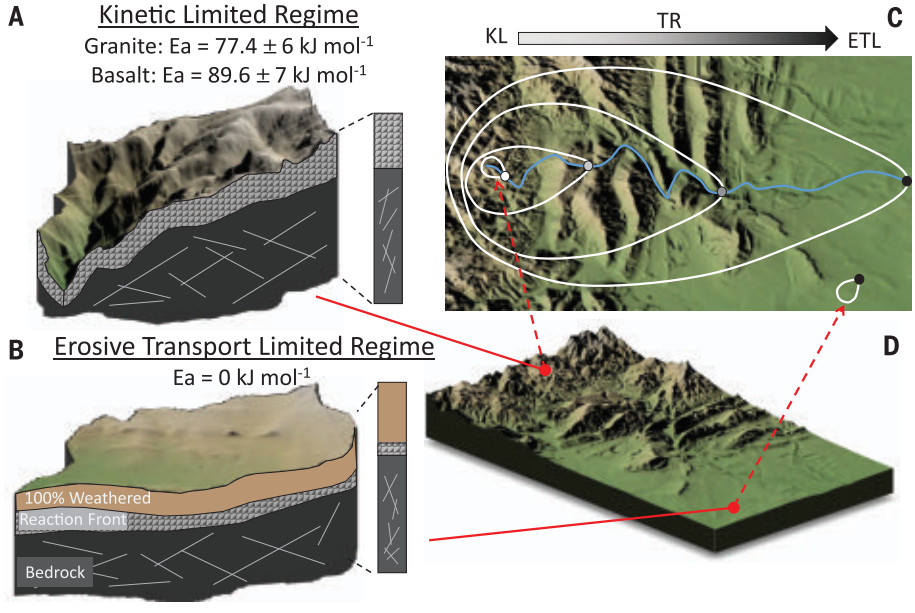


Fig. 1. Conceptual models for soils and watersheds on silicate lithologies in the humid half of Earth's land surface. (**A** to **D**) Reaction fronts [gray hachured layers in (A) and (B)] indicate where most CO_2 is solubilized during silicate weathering. For a bedrock mineral such as $CaSiO_3$, for example, weathering ($2CO_2 + H_2O + CaSiO_3 \rightarrow Ca^{2+} + 2HCO_3^- + SiO_{2(aq)}$) releases dissolved inorganic carbon that is transported, precipitated, and partially sequestered as buried carbonate minerals at the seafloor ($Ca^{2+} + 2HCO_3^- \rightarrow CO_2\uparrow + H_2O + CaCO_{3(s)}$) offshore from the landscapes in (C) and (D). In (A), kinetic limitation (KL) of silicate weathering is indicated by soil profiles or small head catchments where the silicate and its reaction front are exposed everywhere at the land surface, allowing dissolution to proceed far from equilibrium. In (B), weathering of soil profiles and watersheds characterized by buried reaction fronts for bedrock minerals instead can be dominated by flow paths where concentrations of weathering products increase until they cause dissolution to stop. Weathering fluxes from such watersheds can correlate with erosion rate, that is, they are erosive transport–limited (ETL). As shown in (C), in watersheds of increasing size (schematically shown as white teardrop shapes), the decreasing kinetic-limited landscape fraction and increasing dominance of long flow paths drive weathering fluxes to vary with erosion rate (erosive transport–limited). Only small watersheds tend to be kinetic-limited as in (A), but both small (B) and large watersheds [(C) and (D)] can be erosive transport–limited because feedbacks couple weathering fluxes to erosion rates. Thus, as watersheds become larger, they show transition regime (TR) behavior between kinetic- and erosive transport–limited landscapes.

is affected by erosion, temperature, and runoff [e.g., (*14*, *17*)], and the temperature dependences for soils and small and large watersheds do not match (*4*, *14*, *17*–*20*). Additional puzzles include how stream chemistry [e.g., (*21*)] and mountain-building events at different time scales are linked with climate, including how to explain the "Cenozoic isotope-weathering paradox" [e.g., (*6*, *16*, *22*)]. To reconcile observations, we present conceptual models at different spatial and temporal scales and emphasize feedbacks related to how the geological subsurface is structured by weathering reactions and fracturing to guide and respond to groundwater flow (*23*–*25*).

Using standard kinetic treatments, we use the Arrhenius equation (eq. S1) to express temperature sensitivity. This allows calculation of the Arrhenius activation energy, *Ea*, from the natural log of the relative increase in rate as temperature changes from a reference temperature, T_0, to a higher value, T (eq. S2). By using *Ea*, we rely on standard treatments of chemical kinetics by attempting to isolate chemical dissolution from other factors. In the "Global integration" section, we recast *Ea* as the percent change in weathering flux per degree temperature change globally. In contrast to previous researchers, we neither require that weathering is kinetic-limited everywhere nor require that *Ea* is the same at every spatial scale [e.g., (*10*, *26*, *27*)]. Instead, we assume that weathering systems can be treated with temperature-independent *Ea* values at each spatial scale. Although *Ea* might also change with temperature [e.g., (*28*)], our simplified treatment allows parsimonious upscaling to a global model. We discovered that *Ea* varies as additional processes become coupled to dissolution at higher scale. We also emphasize that kinetics can only limit the weathering of landscapes where the exposure time of rocks to weathering is short compared with the reaction time scale [i.e., (*18*)] and the net water balance is positive, that is, mean annual precipitation (MAP) is greater than potential evapotranspiration (PET) [i.e., (*29*, *30*)]. With this approach, the temperature sensitivity of weathering can be reconciled across scales, yielding an approach to understand weathering in the past or in a human-manipulated future.

Analyzing Earth's thermostat

Some authors have explored controls on the long-term carbon balance by quantitatively upscaling based on exposed lithologies (*31*, *32*). For example, global continental land area is ~63% sediments and ~35% crystalline rock, where the latter are ~15% basaltic and ~85% granitic (*31*). A more high-resolution assessment yielded 56.24% of exorheic land area as sediments (excluding carbonates), 24.3% crystalline rock (excluding volcanics), and 7.00% volcanics (86% basaltic) (*32*).

Table 1. Thermodynamic values and activation energies for kinetic-limited systems. $\Delta G°$, standard Gibbs Free Energy.

System	$\Delta G°$ and *Ea* values (kJ mol^{-1})
Granite	
Albite → kaolinite in laboratory	$\Delta G° = 83.5 \pm 2$
Oligoclase → kaolinite in laboratory	$\Delta G° = 74$
Rate-limiting step is dissolution (laboratory)	
Far-from-equilibrium mineral dissolution (oligoclase)*	$Ea_{dissolution} = 67 \pm 2$
Far-from-equilibrium Na release, rock dissolution†	$Ea_{dissolution} = 57 \pm 15$
Coupled dissolution, transport, and/or clay precipitation	
Kinetic-limited soils‡	$Ea_{dissolution} = 79.8 \pm 8$
Coupled dissolution, transport, clay precipitation, fracturing, and/or erosion	
Kinetic-limited watersheds (corrected for MAP and erosion as needed)§	$Ea_{apparent} = 77.4 \pm 6$ (11.2% K^{-1})¶
Transition-regime watersheds (corrected for MAP but not erosion)#	$Ea_{apparent} = 36.1 \pm 15.2$
Theoretical prediction (eqs. S5 and S11)	
$Ea_{dissolution} < Ea_{apparent} < Ea_{diffusion} + Ea_{dissolution} + \Delta G°/3$	$Ea_{apparent} = 67$ to 110
Basalt	
Labradorite → kaolinite in laboratory	$\Delta G° = 64 \pm 7$
Rate-limiting step is dissolution (laboratory)	
Far-from-equilibrium mineral dissolution (labradorite)**	$Ea_{dissolution} = 43 \pm 1$
Si release, crystalline or glassy basalt dissolution††	$Ea_{dissolution} = 46 \pm 2$
Coupled dissolution, transport, and/or clay precipitation	
Olivine in a soil column in laboratory‡‡	$Ea_{apparent} = 72$
Solute release during rind formation on basalts	$Ea_{apparent} = 54 \pm 12$
Solute release during rind formation on andesites and basalts	$Ea_{apparent} = 64 \pm 12$
Coupled dissolution, transport, clay precipitation, fracturing, and/or erosion	
Kinetic-limited watersheds§§	$Ea_{apparent} = 89.6 \pm 7$ (13% K^{-1})¶
Transition-regime watersheds (not corrected for MAP or erosion)#	$Ea_{apparent} = 42.5 \pm 1.3$¶
Theoretical predictions (see supplementary materials)	
Rinds: $Ea_{apparent} = Ea_{diffusion} + \Delta G°/1.6$	$Ea_{apparent} = 55 \pm 7$
Soils: $Ea_{dissolution} < Ea_{apparent} < Ea_{diffusion} + Ea_{dissolution} + \Delta G°/1.6$	$Ea_{apparent} = 43$ to 98

*Weighted average of synthesized laboratory estimates of $Ea_{dissolution}$ for oligoclase in laboratory from table S1A. †As summarized in table S1A (*14*). ‡From values in table S1B, as described in the text. §From values in table S1C, as described in the text. ¶The fractional rate of change with temperature, α, for a given *Ea* approximately equals Ea/RT_0^2, where R is the universal gas constant and T is 288 K, the average Earth surface temperature (*2*, *3*). #Calculated for granitoids (*4*) and the average of basalt fluxes (*61*, *179*). **From table S2A. ††Weighted average Si release rates from crystalline and glassy basalt rock in table S2A (uncorrected for Al). ‡‡See table S2B. §§From Icelandic rivers, as summarized in table S2B. Error is assumed to be ±1% K^{-1}.

We instead analyze a few important minerals that were chosen from those lithologies. Silicate minerals solubilize CO_2 from the atmosphere during weathering and release charged DIC species as the minerals release aqueous base metal cations (Na^+, K^+, Mg^{2+}, Ca^{2+}) with uncharged aqueous silica ($SiO_{2(aq)}$). Ca and Mg are reprecipitated as carbonates in the oceans, thereby sequestering CO_2 and balancing the CO_2 flux from volcanic degassing (*1*, *16*) when considered at time scales greater than 10^5 to 10^6 years. By contrast, K^+ delivered to the ocean may be removed through back reactions with silicates at the seafloor, rereleasing CO_2 to the atmosphere (*33*). K^+ is also immobilized in soils as clays or taken up into biota in soils. A large fraction of the dissolved riverine Na^+ is exchanged with Ca^{2+} and Mg^{2+} on suspended sediments during transport to the oceans (*34*), and these cations in turn sequester CO_2 as (Ca,Mg)-containing carbonates at the seafloor, in effect sequestering DIC associated with Na release as carbonate sediment.

Using these observations, we review previous determinations of *Ea* from plots of log (reaction rate) versus T^{-1} (eqs. S1 and S2) for weathering of the common Ca-, Mg-, and Na-containing silicates that dominate basalts and granites. These include the Ca- and Mg-containing olivine and diopside minerals in basalts and Na- and Ca-containing plagioclase feldspar that is found in both basalts and granites. Plagioclase is particularly important not only because it is the most abundant

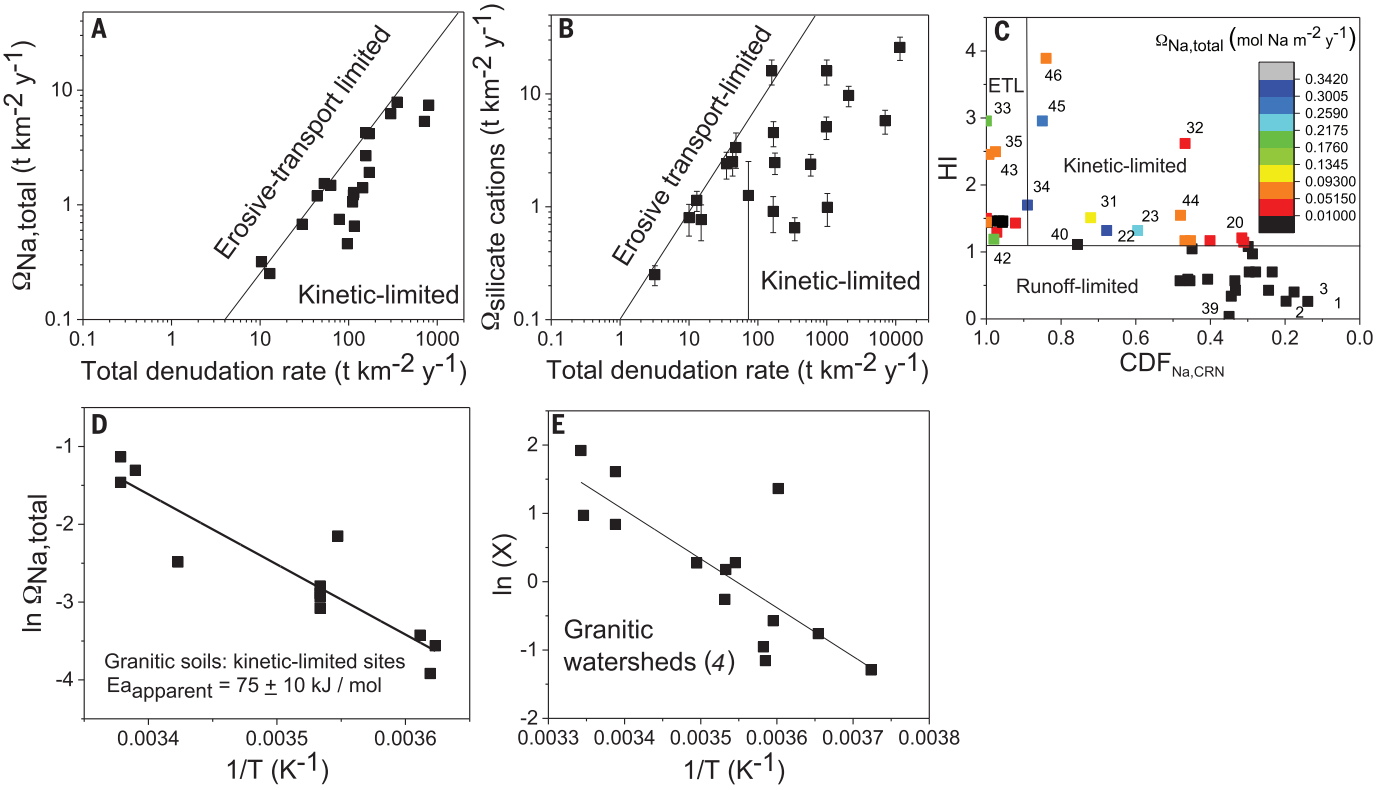


Fig. 2. Weathering rates plotted versus controlling variables. (**A**) Total Na weathering rate, $\Omega_{Na,total}$, for the granitic soils that are not runoff-limited (table S3). (**B**) Total silicate cation weathering rates, $\Omega_{cations,total}$, for erosive transport– and kinetic-limited watersheds from an earlier publication (*4*). In (A) and (B), symbols plotting on and off the line are interpreted as limited by erosive transport or kinetics, respectively, and error bars indicate estimates from the original publication (*4*). (**C**) Summary plot of all 45 soils in table S3 color coded for $\Omega_{Na,total}$ and categorized by HI and $CDF_{Na,CRN}$ (see table S3 for index numbers where indicated). The solid horizontal and vertical lines represent estimated values of HI and CDF that separate the labeled weathering regimes as discussed in the text. (**D** and **E**) Arrhenius plots for the kinetic-limited soils (table S3) (D) and watersheds (E) (*4*). *X* is a weathering rate function corrected for MAP and denudation rate by using parameters from the original authors who calculated $Ea_{apparent}$ = 74 kJ mol^{-1}. Table S3 includes sites with slopes <7.5° from an earlier publication (*18*) after reanalysis (*29*) as well as additional sites from published reports.

crustal mineral but also because it is the source of almost all Na released naturally to rivers, making the assessment of rates of plagioclase weathering from riverine chemistry more robust than that for other abundant rock-forming silicates. This assessment is further simplified because Na usually does not accumulate in biota or clays (an observation revisited in our treatment). Plagioclase feldspar is also one of the more abundant of the reactive, noncarbonate, base metal cation–releasing silicates in the sediments that cover about two-thirds of the continents (*32*).

Laboratory measurements

Comparisons of laboratory kinetics (tables S1A and S2A) reveal that activation energies of dissolution, $Ea_{dissolution}$, are highest for the Mg-releasing minerals (58 to 96 kJ mol^{-1}). For feldspars, values of $Ea_{dissolution}$ are lower and decrease with increasing Ca content from granitic to basaltic compositions. For example, oligoclase ($Ea_{dissolution}$ = 67 ± 2 kJ mol^{-1}) and labradorite ($Ea_{dissolution}$ = 42.8 ± 1 kJ mol^{-1}) are the most common feldspars in granite and basalt, respectively (Table 1).

The compilations also allow comparison of $Ea_{dissolution}$ from isolated-mineral experiments to that of the same mineral dissolved in reactors packed with other phases. For example, the $Ea_{dissolution}$ estimates for Mg-rich olivine (65.6 to 85.7 kJ mol^{-1}) from single-mineral experiments bracket $Ea_{apparent}$ for olivine dissolved in a soil column (72.5 kJ mol^{-1}) (table S2A). We call the latter an $Ea_{apparent}$ because additional steps can enter or couple with dissolution at higher scales. This can change the rate-limiting step or add an additional temperature dependence (*35*).

Similarly, the value of $Ea_{dissolution}$ for a granitic feldspar (oligoclase, 67 ± 2 kJ mol^{-1}) is within error of the $Ea_{apparent}$ that describes Na release from granitic whole-rock experiments, 57 ± 15 kJ mol^{-1} (Table 1 and table S1A), which is expected given that the dominant base metal cation released from dissolving granite is Na. Similarly, although solute release from crystalline basalt is dominated by Mg from olivine at lower pH and Ca from feldspar at higher pH (*36*), we show that the $Ea_{apparent}$ for basalt dissolution is similar to that of feldspar dissolution (table S2A). Specifically, the $Ea_{apparent}$ for Si release over a range of pH from glassy or crystalline basalt, 46 ± 2 kJ mol^{-1}, is within error of the $Ea_{dissolution}$ of labradorite (43 ± 1 kJ mol^{-1}) (Table 1 and table S2A). Given these observations, we emphasize plagioclase feldspar weathering and Na release in the next sections, which treat higher spatial scales.

Soil analysis

Granitic systems

To estimate $Ea_{apparent}$ in natural systems, Na release rates, Ω_{Na}, are assessed for sets of soils known as climosequences that span gradients in mean annual temperature (MAT) [e.g., (*18*)]. For example, a master set of 45 soils from sites such as terraces and ridgetops is summarized in table S3. These soils were chosen because (i) they generally formed on relatively homogeneous parent materials with minimal disturbance by humans or glaciers, and (ii) their chemical and physical denudation rates have been estimated. Ridgetops and terraces are considered the most straightforward to model because they are functionally one-dimensional (1D) in that water and sediments are not transported into them from upslope.

Key to the analysis is that dissolution of feldspar in a 1D soil is localized into a reaction front (e.g., Fig. 1) or depth interval where all reaction occurs [e.g., (*37*)]. Below the front, pore fluids generally flow more slowly and are sufficiently concentrated so that they react only minimally with feldspar. At any depth in such a 1D profile, the fraction of mineral depleted by weathering can be estimated compared with the parent rock concentration. This depth-dependent fraction, termed the chemical depletion fraction (CDF; estimated from eq. S4), varies from zero in the parent material to values ≤100% above or at the top of the front. The error in CDF is larger in soils when CDF is small and in regions of substantial dust additions or parent composition variability.

Profiles with 0 < CDF < 100% for a dissolving mineral at the land surface (i.e., Fig. 1A) are "incompletely developed" because they retain the mineral throughout the weathering column. When such profiles are identified for plagioclase feldspar, exposure times (for non-eroding soils) or residence times (for eroding soils) are inferred to be short compared with the time scale of plagioclase dissolution (*37*). Instead of concentrations of plagioclase, however, Na is measured as a proxy because most Na in granitic material is typically in feldspar. We also assume for eroding 1D soils that the erosion rate and weathering advance rate are roughly balanced (expressed in units of length per time), which explains why soils endure and can remain roughly constant in thickness over the soil residence time. With those assumptions, mass balance arguments show that CDF values calculated at the land surface for Na, $CDF_{Na,surf}$, equal the fraction of the total denudation rate (D, in tonnes km^{-2} $year^{-1}$) of Na caused by chemical weathering (*18*). For example, $CDF_{Na,surf}$ = 100% implies 100% chemical removal of plagioclase and negligible loss to erosion (Fig. 1A). The implication from models is that incomplete (0 < $CDF_{Na,surf}$ < 100%) or complete development of a 1D profile ($CDF_{Na,surf}$ = 100%) reflects kinetic or erosive-transport limitation of plagioclase weathering, respectively (*37*).

For most soils in table S3, weathering release rates of Na were calculated from $CDF_{Na,surf}$ by following approaches based on cosmogenic radionuclide (CRN) concentrations (*18*). This approach yields Na release rates, $\Omega_{Na,CRN}$, calculated as the product of the denudation rate, D_{CRN}, multiplied by $CDF_{Na,CRN}$ and $[Na]_{rock}$ (*18*). The subscript CRN is used because $CDF_{Na,CRN}$ and D_{CRN} are the $CDF_{Na,surf}$ and denudation rate, respectively, averaged over the surficial layer of soil where penetration of cosmic rays result in accumulation of CRN. $[Na]_{rock}$ is the Na concentration of bedrock. Later treatments (*29*) corrected these rates to include weathering in saprolite below the layer of CRN accumulation (*38*). We refer to these corrected weathering and denudation rates for Na as $\Omega_{Na,total}$ and $D_{Na,total}$, respectively (see table S3).

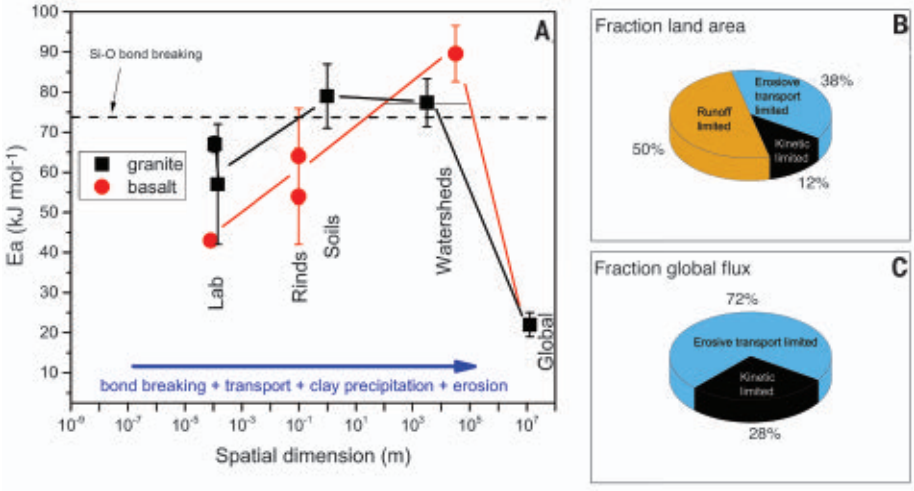


Fig. 3. Global assessment of *Ea* across spatial scales. (**A** to **C**) Variation in *Ea* for kinetic-limited weathering from laboratory to rind formation to soils to watersheds. Weathering mechanisms include additional steps as the spatial dimension increases from ~100-μm mineral grains to ~10-cm clasts to ~1-m soils to ~10^2- to 10^5-m drainage basins. The calculated *Ea* of bond breaking of an edge site for tetrahedral silicate anions is shown as 74 kJ mol^{-1}. The calculated *Ea* for silicate weathering for the global Earth was not upscaled based on fractional land areas of weathering regimes (B) but instead on fractional contributions of fluxes (C). First, a temperature response was estimated based on the relative proportions of global CO_2 consumption fluxes from the two main lithologies, granite (74%) and basalt (26%) (11.7% K^{-1}). This was then used to upscale *Ea* to the globe based on the proportions from each weathering regime (C) where the contribution from runoff-limited watersheds was assumed to be negligible. Symbols in (A) show averages with 1σ from Table 1, where the laboratory symbols represent oligoclase or rock (granite) and labradorite (basalt); the two rind symbols represent basaltic and andesitic-basaltic compositions. See text for further explanations.

To calculate $Ea_{apparent}$ for weathering, we focused on 1D granitic soils that are incompletely developed with respect to Na because such soils are often kinetic-limited, and that weathering regime is the only one that we expect to show strong temperature sensitivity (*18*). Completely developed soils (90% ≤ $CDF_{Na,CRN}$ ≤ 100%) were not considered on the assumption that they are limited by erosion rate rather than kinetics and therefore are not strongly dependent on MAP or MAT (fig. S3): Such fluxes for soils listed in table S3B plot on the line in Fig. 2A. The threshold CDF for completely developed profiles was dropped from 100 to 90% because $CDF_{Na,CRN}$ generally is only known to about ±10% (*35*). This threshold also allowed identification of one soil as erosive transport–limited rather than kinetic-limited, which is consistent with its extreme weathering depth (table S3B).

We also did not consider soils with incompletely developed reaction fronts for feldspar in calculations of $Ea_{apparent}$ when they developed under a negative local water budget. For such soils, weathering appears limited by lack of water flushing instead of by weathering kinetics (*29*). Na reaction fronts in some of these soils extend to great depths, but Na is nonetheless retained at the land surface. This occurs in some cases because of the precipitation of Na-containing smectitic clays (*39*). We refer to this dry regime as limitation by runoff, where runoff is equated to MAP minus PET.

To identify runoff limitation, we could not rely on clay compositions or depth of total weathering because these were not always reported for the soils listed in table S3. We followed an earlier approach (*29*) and calculated the approximate threshold for runoff limitation as the point where the humidity index (HI), defined as MAP divided by PET, falls below unity. But errors in MAP and PET make this threshold very approximate: For example, a value of unity did not allow inclusion of a soil with a deep, incompletely developed Na profile (table S3A) where smectite, a geologically long-term indicator of retention of base metal cations (*35*), had been reported. That site, characterized by HI = 1.1, has a relatively high MAP with a very long dry season. To include that soil in the regime of runoff limitation, we set the threshold HI at 1.1.

Clearly, such ad hoc definitions for runoff limitation are potentially problematic, but the small number of soils with $\Omega_{Na,total}$ estimates based on the CRN approach (table S3) precludes rigorous testing. In addition, runoff limitation has not been investigated as a

regime of weathering by most authors, despite the acknowledged importance of runoff [e.g., (*4*, *20*, *29*)]. We therefore further tested the idea of runoff limitation as a function of HI by investigating watershed data. We focused on a commonly measured weathering-derived solute (SiO_2) and assessed silica release rates, Ω_{SiO2}, versus runoff (defined here as stream discharge normalized by basin area). To simplify our analysis to only hypothetically kinetic- and runoff-limited watersheds, we targeted watersheds on highly erodible shale, assuming that few shales would weather in the erosive transport–limited regime (*35*). Validating expectations for runoff and kinetic limitation, respectively, shale-derived Ω_{SiO2} data are strongly runoff-dependent at low HI and highly temperature-dependent at high HI (fig. S2, D and F). Given that Ω_{SiO2} is determined from concentration (C) times discharge (q) normalized by basin area (*35*), a linear dependence of Ω_{SiO2} on runoff (fig. S2E) implies that C is independent of q (fig. S2H). Figure S2H highlights the possibility that such discharge-independent or "chemostatic" (*21*) concentrations in low-HI watersheds could reflect smectite-kaolinite equilibration. This reaction could also explain why low-HI SiO_2 fluxes do not generally increase with temperature (fig. S2F). Specifically, the stoichiometry and negative enthalpy of this reaction require that concentrations of SiO_2 decrease as the equilibrium shifts toward smectite with increasing temperature.

To determine the temperature sensitivity of weathering for the 45-soil dataset, we therefore eliminated both runoff- and erosive-transport limited sites by using HI and CDF criteria, respectively, and focused on the 12 remaining kinetic-limited soils (table S3C). We also considered temperature correlations for MAP and $D_{Na,total}$ for the 12 soils. Such correlations occur because precipitation increases with temperature over broad spatial scales [e.g., (*40*)] and because chemical weathering is coupled to physical erosion (*18*). We used a power-law treatment to explore if $\Omega_{Na,total}$ should be corrected for such correlations. That treatment (eq. S3c) had previously been used to correct values of $\Omega_{Na,CRN}$ for MAP and D_{CRN} (*18*), yielding an $Ea_{apparent}$ equal to 17 kJ mol^{-1}, which is much lower than the $Ea_{dissolution}$ for plagioclase (Table 1). However, with that "correction," much of the temperature dependence in $\Omega_{Na,CRN}$ was effectively attributed to denudation because, counter to expectations for kinetic limitation, $CDF_{Na,CRN}$ shows little variation with denudation. We similarly observed such a lack of variation in $CDF_{Na,CRN}$ with $D_{Na,total}$ (fig. S4A), but we embraced the alternate interpretation posed by the researchers (*18*), namely, that $\Omega_{Na,CRN}$ correlates with denudation because erodibility increases as chemical weathering weakens and disaggregates rock [e.g., (*41*, *42*)]. We also argue (*35*) that error in CDF obscures variations with denudation in some but not all fast-eroding systems (*18*, *43*).

Instead of following the earlier approach (*18*), we therefore attributed the temperature dependence of denudation to the $Ea_{apparent}$ of weathering and its effect on rock disintegration into blocks or particles. Without using the power-law correction, $Ea_{apparent}$ for the 12 soils in table S3C equals 75 ± 10 kJ mol^{-1} (Fig. 2D). Corroborating this value are three other estimates between 75 ± 14 and 112 ± 34 kJ mol^{-1} from two additional kinetic-limited climosequences on loess and shale (table S1B), where measurements of concentrations to depths of meters removes some but not all ambiguities in CDF_{Na} (*35*). Corrections for MAP and $D_{Na,total}$ are considered irrelevant on the loess and shale because MAP and $D_{Na,total}$ vary negligibly or are temperature-independent. The weighted average of all the estimates, 79.8 ± 8 kJ mol^{-1} (Table 1), is considered our best estimate of granitic soil–based $Ea_{apparent}$.

Basaltic systems

No soil-based estimates of Ea describe the weathering of basaltic soils based on CRNs. Instead, 13 reports of weathering rind formation on basalts or the less Mg-rich andesites were found with HI > 1.1 along with a few measurements of pitting on naturally exposed basaltic feldspars (tables S2B, S4, and S5). For these, rind formation initiates as plagioclase alters to clay. For HI > 1.1, the best estimate of $Ea_{apparent}$ for rind formation equals 54 ± 12 kJ mol^{-1} for basalts and 64 ± 12 kJ mol^{-1} for basalts and andesites (*35*). These values overlap with the range of $Ea_{apparent}$ that was estimated for one study of porosity development on naturally exposed basaltic plagioclase, 56 ± 16 to 97 ± 10 kJ mol^{-1} (table S2B).

Watershed analysis

Granitic systems

In this section, we analyze watershed studies that report $Ea_{apparent}$. The published studies for granitic watersheds (table S1C) focus on Na, Si, or base metal–cation release rates corrected for nonsilicate-derived cations. As expected for studies that mostly emphasize permanent streams, almost all $Ea_{apparent}$ estimates are for systems at relatively high HI. This partly explains why runoff limitation has not been defined previously for watersheds [however, see (*4*, *20*, *21*)]. Separation of runoff-limited watersheds is nonetheless important because only the analysis of high-HI watersheds yields an $Ea_{apparent}$ value that is significant (fig. S2F).

To calculate a best estimate of $Ea_{apparent}$ for watersheds to compare with that of other spatial scales, we focused on Na or carbonate-corrected silicate cation release rates ($\Omega_{Na,total}$ or $\Omega_{cations,total}$, respectively). Treating these fluxes together relies on the assumption that their $Ea_{apparent}$ values are comparable because feldspars dominate granitic cation release. We did not include Ω_{SiO2} because Si-based $Ea_{apparent}$ is consistently lower for watersheds where release rates of cations and Si are both reported (table S1C). As for soils, we also considered potential temperature correlations for MAP and D. In some cases, these correlations were not observed. For example, in one compiled dataset from >40 head catchments, values of $\Omega_{Na,total}$ were corrected for MAP but not denudation because $\Omega_{Na,total}$ showed no variation with a proxy for erosion (relief) (*44*). Three other studies targeted only a few neighboring watersheds, negating the need for regional corrections of MAP or denudation. The fifth study (see Fig. 2, B and E) corrected $\Omega_{cations,total}$ for 13 alpine or montane watersheds for both MAP and D_{total} by fitting $\Omega_{cations,total}$ to a version of power-law eq. S3c, yielding $Ea_{apparent}$ = 74 ± 29 kJ mol^{-1} (*4*). The weighted average of these watershed $Ea_{apparent}$ values for granites equals 77.4 ± 6 kJ mol^{-1}.

Before accepting this estimate, however, we address an apparent contradiction. Specifically, we used watershed values of Ω that were corrected for MAP and D (*4*) by fitting to the same power-law equation (*18*) that we rejected for soils. The fit [see also (*45*)] showed that the chemical weathering correlation with total denudation is weaker for watersheds (power-law exponent $a < 1$) than soils ($a \approx 1$) (*4*). But in addition, the fits differ in that the correction for D for soils lowers the $Ea_{apparent}$ (compared with its uncorrected value), whereas the correction for watersheds increases $Ea_{apparent}$. We explore this difference next using Fig. 1.

In Fig. 1A, we propose that kinetic-limited watersheds are like kinetic-limited soils in that they retain feldspar at the surface. However, in watersheds, the soils are not the only factor that dictates stream chemistry. Conceptually, release rates of silicate-derived cations such as Na are kinetic-limited only in watersheds that are small enough that ground waters flowing from soil to stream cumulatively interact with little feldspar. Short flow paths and residence times preclude the attainment of chemical equilibrium [see, for example, (*46*)]. The result is that long-term average stream concentrations in kinetic-limited watersheds vary with discharge because each flow path is limited by temperature-dependent kinetics ($Ea_{apparent} \approx 80$ kJ mol^{-1}) and not by equilibrium.

The other type of head catchment is characterized by feldspar-absent topsoils (Fig. 1B) and stream concentrations that are more likely to be roughly constant (chemostatic) as a function of variable discharge when observed over years to decades. Because water chemistry is unlikely to equilibrate with feldspar at ambient conditions even for long flow paths (*47*), this

chemostasis could reflect a limiting reagent in the dissolving waters, steady-state concentrations that are dictated by the dissolution of primary minerals balanced by precipitation of secondary minerals, or equilibration among secondary minerals [e.g., (*5*, *21*, *48*, *49*)]. But for small granitic watersheds similar to that shown in Fig. 1B, it has been observed that their weathering fluxes vary strongly with erosion rate rather than temperature [e.g., (*15*)]. For this reason, we label them erosive transport–limited and assume $Ea_{\text{apparent}} \approx 0$ kJ mol^{-1}.

As watersheds become larger (Fig. 1C), they are increasingly dominated by lowlands [e.g., (*20*)] and by chemostatic behavior because groundwater flow paths become deeper, longer, and more likely to interact with feldspar for extended residence times [e.g., (*46*)]. But large watersheds are similar to the watersheds in Fig. 1B in that their weathering fluxes are observed to depend strongly on erosion rate and not temperature [e.g., (*20*)]. The fluxes of these large watersheds are thus considered erosive transport–limited: The average product of flow-weighted C and q depends on erosion rate (*20*).

Returning to the power-law corrections for Ω, we argue that fitting watersheds using eq. S3c adjusts for the fraction of soil landscapes and flow paths that are kinetic-limited ($Ea_{\text{apparent}} \approx 80$ kJ mol^{-1}) versus erosive transport–limited (~0 kJ mol^{-1}). Although the large watersheds may contain kinetic-limited headwater catchments, these do not dominate mainstem chemistry downstream. The power-law correction to calculate temperature sensitivity of weathering thus yields a higher, rather than lower, Ea_{apparent}. Two corollaries to this argument can be tested. First, if this argument is defensible, then the maximum weathering fluxes for aggregated datasets for soils and watersheds should be similar. This has been documented (*4*, *50*) (see also Fig. 2). The second corollary is that virtually all kinetic-limited watersheds should be small, whereas erosive transport–limited watersheds should be both large and small. This has also been observed (*4*, *50*).

But, why should averaged fluxes (= $C{\cdot}q$) from erosive transport–limited watersheds tend toward values dictated by erosion rate? The most likely answer is feedbacks that couple water chemistry, subsurface flow paths, fracturing, and erosion at the watershed scale. For example, permeability changes driven by weathering and fracturing [e.g., (*23*, *51*)] could comprise feedbacks that couple weathering flux to erosion. Some aspects of such feedbacks have been explored for pyrite weathering (*24*, *25*). Processes related to vegetation are also potential feedbacks [e.g., (*52*)].

Basaltic systems

The reported values of Ea_{apparent} for basaltic watersheds ranged as high as 89.6 kJ mol^{-1} (table S2B). That value derives from an Ea_{apparent} for Icelandic volcanic watersheds after correction for runoff and glacial melting (*53*). Conversely, the Icelandic fluxes were not corrected for issues related to carbonates formed from volcanic CO_2 (*54*). Those types of complications, including deep subsurface weathering and volcanic degassing [e.g., (*55*)], led one set of researchers to estimate global CO_2 consumption for basalts by focusing only on 22 inactive volcanic systems (table S2B). Their calculated Ea_{apparent}, which was much lower than the Icelandic value, was similar in magnitude to an earlier estimate based on 10 systems (table S2B). No evidence requiring a correction for runoff, MAP, or denudation rate was reported in either study. Weighting and averaging the two lower values yields 42.5 ± 1.3 kJ mol^{-1} (table S2B).

To reconcile these high and low values, we propose, by analogy to granitic watersheds, that the higher Ea_{apparent}, 89.6 ± 7 kJ mol^{-1}, for the Icelandic systems (*19*, *53*) is the best estimate for kinetic-limited basaltic landscapes and the lower value, 42.5 ± 1.3 kJ mol^{-1}, reflects contributions from both kinetic- and erosive transport–limited soils within transition-regime watersheds (Fig. 1). As summarized in table S2B, the higher Ea_{apparent} corroborates previous treatments of cross-scale weathering that yielded 70 ± 20 kJ mol^{-1} (*26*) and seafloor basalt weathering based on inverse modeling that yielded 75 ± 22 to 92 ± 7 kJ mol^{-1} (*6*, *12*). Likewise, values of Ea_{apparent} that range as high as 109 kJ mol^{-1} have been reported from the pitting of naturally exposed plagioclase grains (table S2B).

Cross-scale comparison

Many additional lines of evidence confirm aspects of these interpretations. For example, molecular calculations of Ea of hydrolysis of Si tetrahedral networks at edges or corners on Na feldspar surfaces, 74 to 120 kJ mol^{-1} (*56*), are slightly higher than the laboratory $Ea_{\text{dissolution}}$ values (43 to 67 kJ mol^{-1}), probably because of proton sorption (*57*). If we invoke arguments related to differences in Al content (*56*), molecular calculations are also consistent with the laboratory-based conclusion that $Ea_{\text{dissolution}}$ values for the dominant granitic feldspars are higher (67 kJ mol^{-1}) than those for basaltic feldspars (43 kJ mol^{-1}).

The laboratory values of $Ea_{\text{dissolution}}$ for the Na- and Ca-feldspars are within error of the values for granite and basalt whole-rock dissolution, respectively, but the Ea_{apparent} values increase by at least 10 kJ mol^{-1}or as much as 50 kJ mol^{-1} for soils formed on granite and for watersheds formed on basalts, respectively (Table 1 and Fig. 3). Biota might cause this lab-field difference because we know very little about the biotic effects on Ea (tables S1 and S2). However, where biotic effects have been measured, they sometimes decrease rather than increase the Ea_{apparent} (tables S1A and S2, A and B), and neither biota nor types of land cover are reported in recent studies as major factors that control weathering rates in watersheds (*7*, *58*). The biotic effect may be small and largely caused by pH because organic acids are low in concentration in soils (*59*). We averaged Ea values across pH ranges where measurements were made.

Instead, we argue that this ~10 kJ mol^{-1} (granite) or 50 kJ mol^{-1} (basalt) laboratory-field discrepancy arises because feldspar weathering in the laboratory is studied far from equilibrium and depends strictly on dissolution, whereas the mechanism in natural systems includes contributions from processes that enter at higher spatial scales and that cannot be entirely separated, including transport, precipitation of clays, rock disaggregation, and fracturing (*25*). For example, theoretical predictions for dissolution-precipitation coupling (eqs. S11, S17, and S18) may explain the value for both granites and basalts (Table 1). Likewise, researchers have argued that the high Ea_{apparent} for Icelandic basalts reflects dissolution and secondary mineral precipitation (*53*). The inextricable coupling of erosion and chemical weathering [e.g., (*15*, *18*, *20*)] also points to temperature-sensitive processes such as disaggregation or subsurface flow partitioning that contribute to soil- or watershed-based Ea_{apparent} values.

Global integration

Assuming that these approaches to treating chemical weathering, physical denudation, and hydrology are defensible, we now upscale to the global system by calculating the fractional change in silicate weathering flux with each degree increase in temperature, α_{sw}, for the kinetic-limited regime using the approximation that $\alpha_{sw} \approx \frac{Ea}{RT_0^2}$ (*2*, *3*), where R is the universal gas constant. After setting T_0 to the global mean temperature of 288 K, the values of α_{sw} equal 11.2 and 13% K^{-1} for kinetic-limited granitic and basaltic catchments (Table 1), respectively.

To upscale requires assessing α_{sw} for global distributions of lithology and weathering regime. Instead of upscaling for lithology by area, however, we instead upscaled using proportions of CO_2 consumption flux as a function of lithology. This strategy reflects that the fractional areas for each lithology do not translate directly to the proportion of flux because of uneven geographic distributions of lithology and MAP [e.g., (*32*, *60*)]. To upscale, we note that the CO_2 consumption flux from basaltic weathering is estimated based on rivers as 16.5 to 35% (average = 26%) of the total global silicate CO_2 consumption flux (*20*, *32*, *61*). Based on this, we calculated the lithologic flux-weighted α_{sw} of 11.7% K^{-1} from the sum of the granitic

[(100% – 26%) × (11.2% K^{-1})] and basaltic terms [26% × (13% K^{-1})].

A second aspect of the upscaling is assessing the proportions of fluxes for soils in the three weathering regimes. A recent analysis (*30*) suggests that the fraction of global landscape with an HI less than the threshold we defined for runoff limitation is 0.496 (*35*). This means that half the land surface is wet and half dry or that half is fast-weathering and half is slow-weathering [see, for example, figure 6 of (*32*)]. To subdivide the fast-weathering wet half into kinetic- and erosive transport–limited fractions, we note that only about 9 to 14% (average ≈ 12%) of the exorheic land surface contributes 38 to 50% (average = 44 ± 6%) of the total global chemical denudation (*32*, *45*, *62*). This global CO_2 consumption flux from ~12% of the continental area is considered here to represent kinetic limitation because such watersheds tend to weather at the fastest rates (e.g., Fig. 2C) (*50*). However, only 63% of the global flux is from silicates, whereas the rest is from carbonates (*32*). Therefore, although the global fractions of land areas that are kinetic-, runoff-, and erosive transport–limited are about 12, 50, and 38%, respectively, the global fractions of flux from silicates in each regime are about 28(±4)% [= 0.63 · (0.44 ± 0.06)], 0%, and 72% respectively (Fig. 3, B and C). A simple estimate for $\frac{Ea}{RT_0^2}$ for global CO_2 consumption flux from silicate weathering based on the global lithology-weighted average α_{sw} of 11.7% therefore equals ~3.2% K^{-1} [= 28(±4)% · 0.117]. Allowing the fraction of kinetic-limited flux to vary from 38 to 50% yields an $Ea_{apparent}$ value of 22 ± 3 kJ mol^{-1}. We have assumed that (i) weathering of erosive transport–limited landscape depends only on erosion rate, (ii) the weathering flux from runoff-limited landscape is a negligible fraction of the global CO_2 consumption flux, and (iii) weathering of kinetic-limited landscape depends only on temperature. This latter assumption emphasizes that variations of chemical weathering (Ω) with MAP or runoff can be poorly handled by simple regression equations (e.g., eq. S3), given the threshold behavior we identified for chemical weathering versus HI. In that regard, our estimate is probably a lower limit. For example, if we assume that the flux from the kinetic-limited landscape responds to temperature by 11.7% K^{-1} but also to runoff increases driven by regional rises in temperature at 2.7% K^{-1} (*40*), then the $Ea_{apparent}$ increases to 28 kJ mol^{-1}.

Our model emphasizes that high proportions of the globe are soil-shielded or very dry, where weathering is temperature-insensitive or negligible, respectively. Even though weathering of drylands could increase with increases in MAP from warming, their minimal weathering fluxes preclude large contributions to global weathering. Our estimate of 22 ± 3 kJ mol^{-1} is within the range of values inferred from inverse arguments based on recent global paleoclimate models (table S1C) [e.g., (*12*, *63*)]. Those low sensitivities were required within the paleoreconstruction models to reproduce reasonable atmospheric P_{CO_2} levels. Our model is also consistent with previous assertions about mountain-building episodes, where uplift can simultaneously focus precipitation on and exhume previously runoff-limited or shielded landscape [e.g., (*64*)]. After the India-Asia collision during the Tertiary, for example, the Himalayan uplift not only exposed new mineral surface and source rocks [e.g., (*65*)] to weathering but also redistributed precipitation locally. With such "teleconnections," CO_2 consumption fluxes could have remained near-constant globally, maintaining a balanced carbon budget (*16*). But if newly uplifted landscape was no longer dry or soil-shielded and instead became kinetic-limited, then the net global sensitivity of weathering to temperature could have increased, leading to a more stable climate system [see also (*5*)].

However, our treatment mostly explains some contradictions in the literature while emphasizing the importance of scale and runoff limitation. These factors of scale and runoff highlight a need for paleogeographic reconstructions in assessing long-term temperature sensitivity [e.g., (*60*)]. We also need to consider such factors if we try to enhance weathering for atmospheric CO_2 drawdown. Mining, grinding, and amending global soils with basalt (*66*) might accelerate weathering into the future but will be most successful if water flow is high enough and pathways short enough to remain in the regime of kinetic limitation. More attention should also focus on understanding factors such as pore-water chemistries, clays, particle sizes, fractures, biotic effects, and subsurface flow paths that couple physical and chemical weathering during natural or enhanced weathering.

REFERENCES AND NOTES

1. T. Volk, *Am. J. Sci.* **287**, 763–779 (1987).
2. R. A. Berner, *Am. J. Sci.* **294**, 56–91 (1994).
3. B. P. C. Hayworth, B. J. Foley, *Astrophys. J. Lett.* **902**, L10 (2020).
4. A. J. West, A. Galy, M. Bickle, *Earth Planet. Sci. Lett.* **235**, 211–228 (2005).
5. K. Maher, C. P. Chamberlain, *Science* **343**, 1502–1504 (2014).
6. L. A. Coogan, S. E. Dosso, *Earth Planet. Sci. Lett.* **415**, 38–46 (2015).
7. K. Deng, S. Yang, Y. Guo, *Nat. Commun.* **13**, 1781 (2022).
8. H. C. Urey, *The Planets*, *Their Origin and Development* (Yale Univ. Press, 1952).
9. M. J. Winnick, K. Maher, *Earth Planet. Sci. Lett.* **485**, 111–120 (2018).
10. J. C. G. Walker, P. B. Hays, J. F. Kasting, *J. Geophys. Res.* **86**, 9776–9782 (1981).
11. R. J. Graham, R. Pierrehumbert, *Astrophys. J.* **896**, 115 (2020).
12. J. Krissansen-Totton, D. C. Catling, *Nat. Commun.* **8**, 15423 (2017).
13. P. V. Brady, *J. Geophys. Res.* **96**, 18101–18106 (1991).
14. A. F. White *et al*., *Geochim. Cosmochim. Acta* **63**, 3277–3291 (1999).
15. R. Millot, J. Gaillardet, B. Dupre, C. J. Allegre, *Earth Planet. Sci. Lett.* **196**, 83–98 (2002).
16. L. Kump, M. A. Arthur, in *Tectonic Uplift and Climate Change*, W. F. Ruddiman, Ed. (Springer, 1997), pp. 400–427.
17. Y. Huh, G. Panteleyev, D. Babich, A. Zaitsev, J. M. Edmond, *Geochim. Cosmochim. Acta* **62**, 2053–2075 (1998).
18. C. S. Riebe, J. W. Kirchner, R. C. Finkel, *Earth Planet. Sci. Lett.* **224**, 547–562 (2004).
19. S. R. Gislason *et al*., *Earth Planet. Sci. Lett.* **277**, 213–222 (2009).
20. J. Gaillardet, B. Dupre, P. Louvat, C. J. Allegre, *Chem. Geol.* **159**, 3–30 (1999).
21. S. E. Godsey, J. Hartmann, J. W. Kirchner, *Hydrol. Processes* **33**, 3056–3069 (2019).
22. M. E. Raymo, W. F. Ruddiman, *Nature* **359**, 117–122 (1992).
23. S. L. Brantley *et al*., *Geomorphology* **277**, 100–117 (2017).
24. A. R. Shaughnessy, X. Gu, T. Wen, S. L. Brantley, *Hydrol. Earth Syst. Sci.* **25**, 3397–3409 (2021).
25. X. Gu, P. J. Heaney, F. D. A. A. Reis, S. L. Brantley, *Science* **370**, eabb8092 (2020).
26. A. Navarre-Sitchler, S. L. Brantley, *Earth Planet. Sci. Lett.* **261**, 321–334 (2007).
27. S. A. Banwart, A. Berg, D. J. Beerling, *Global Biogeochem. Cycles* **23**, GB4013 (2009).
28. J. Lloyd, J. A. Taylor, *Funct. Ecol.* **8**, 315–323 (1994).
29. C. Rasmussen *et al*., *Earth Planet. Sci. Lett.* **301**, 521–530 (2011).
30. S. Calabrese, A. Porporato, *Environ. Res. Commun.* **2**, 085005 (2020).
31. P. Amiotte Suchet, J.-L. Probst, W. Ludwig, *Global Biogeochem. Cycles* **17**, 1038 (2003).
32. J. Hartmann, N. Jansen, H. H. Durr, S. Kempe, P. Kohler, *Global Planet. Change* **69**, 185–194 (2009).
33. F. T. Mackenzie, R. M. Garrels, *Am. J. Sci.* **264**, 507–525 (1966).
34. F. L. Sayles, P. C. Mangelsdorf Jr., *Geochim. Cosmochim. Acta* **41**, 951–960 (1977).
35. Materials and methods are available as supplementary materials.
36. S. Gudbrandsson, D. Wolff-Boenisch, S. R. Gislason, E. H. Oelkers, *Geochim. Cosmochim. Acta* **75**, 5496–5509 (2011).
37. M. I. Lebedeva, R. C. Fletcher, S. L. Brantley, *Earth Surf. Process. Landf.* **35**, 508–524 (2010).
38. J. I. Dixon, A. M. Heimsath, R. Amundson, *Earth Surf. Process. Landf.* **34**, 1507–1521 (2009).
39. Y. Tardy, *Chem. Geol.* **7**, 253–271 (1971).
40. M. P. Byrne, P. A. O'Gorman, *J. Clim.* **28**, 8078–8092 (2015).
41. G. Dolui, S. Chatterjee, N. Das Chatterjee, *Model. Earth Syst. Environ.* **2**, 132 (2016).
42. B. P. Murphy, J. P. L. Johnson, N. M. Gasparini, L. S. Sklar, *Nature* **532**, 223–227 (2016).
43. I. J. Larsen *et al*., *Science* **343**, 637–640 (2014).
44. A. F. White, A. E. Blum, *Geochim. Cosmochim. Acta* **59**, 1729–1747 (1995).
45. I. J. Larsen, D. R. Montgomery, H. M. Greenberg, *Geology* **42**, 527–530 (2014).
46. K. Maher, *Earth Planet. Sci. Lett.* **294**, 101–110 (2010).
47. C. Zhu *et al*., *Geochim. Cosmochim. Acta* **271**, 132–153 (2020).
48. L. Li *et al*., *Earths Futur.* **10**, e2021EF002603 (2022).
49. D. E. Ibarra *et al*., *Geochim. Cosmochim. Acta* **190**, 265–293 (2016).
50. J. L. Dixon, F. von Blanckenburg, *C. R. Geosci.* **344**, 597–609 (2012).
51. X. Gu *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 18991–18997 (2020).
52. R. Amundson, A. Heimsath, J. Owen, K. Yoo, W. E. Dietrich, *Geomorphology* **234**, 122–132 (2015).
53. E. S. Eiriksdottir, S. R. Gislason, E. H. Oelkers, *Geochim. Cosmochim. Acta* **107**, 65–81 (2013).
54. M. G. Andrews, A. D. Jacobson, *Geochim. Cosmochim. Acta* **215**, 247–262 (2017).
55. S. Rad, K. Rive, C. J. Allegre, *Aquat. Geochem.* **17**, 221–241 (2011).
56. L. J. Criscenti, J. D. Kubicki, S. L. Brantley, *J. Phys. Chem. A* **110**, 198–206 (2006).
57. S. Nangia, B. J. Garrison, *Mol. Phys.* **107**, 831–843 (2009).
58. F. von Blanckenburg *et al*., *Am. J. Sci.* **321**, 1111–1163 (2021).
59. O. S. Pokrovsky, L. S. Shirokova, S. A. Zabelina, G. Jordan, P. Benezeth, *NPJ Mater. Degrad.* **5**, 51 (2021).
60. Y. Goddéris, Y. Donnadieu, G. Le Hir, V. Lefebvre, E. Nardin, *Earth Sci. Rev.* **128**, 122–138 (2014).

61. C. Dessert, B. Dupre, J. Gaillardet, L. Francois, C. Allegre, *Chem. Geol.* **202**, 257–273 (2003).
62. J. D. Milliman, K. L. Farnsworth, *River Discharge to the Coastal Ocean: A Global Synthesis* (Cambridge Univ. Press, 2011).
63. T. D. Herbert *et al*., *Science* **377**, 116–119 (2022).
64. P. Maffre *et al*., *Earth Planet. Sci. Lett.* **493**, 174–185 (2018).
65. J. K. Willenbring, F. von Blanckenburg, *Nature* **465**, 211–214 (2010).
66. D. J. Beerling *et al*., *Nature* **583**, 242–248 (2020).

ACKNOWLEDGMENTS

We thank L. Kump, J. Kasting, C. Rasmussen, J. Dixon, X. Gu, F. von Blanckenburg, A. Dere, T. Paces, and three anonymous reviewers for contributions. **Funding:** S.L.B. acknowledges support from the Hubert L. Barnes and Mary Barnes endowment funds (Penn State). A.S. acknowledges graduate fellowship support from the US National Science Foundation. **Author contributions:** Conceptualization: S.L.B.; Methodology: S.L.B., A.S., M.I.L., V.N.B.; Investigation: S.L.B., A.S.; Visualization: S.L.B., A.S.; Project administration: S.L.B., A.S.; Writing: S.L.B. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add2922
Materials and Methods
Supplementary Text
Figs. S1 to S4
Tables S1 to S5
References (*67*–*210*)

Submitted 2 June 2022; accepted 22 November 2022
10.1126/science.add2922

NANOPHOTONICS

Collective super- and subradiant dynamics between distant optical quantum emitters

Alexey Tiranov[1]*†, Vasiliki Angelopoulou[1]†, Cornelis Jacobus van Diepen[1]†, Björn Schrinski[1], Oliver August Dall'Alba Sandberg[1], Ying Wang[1], Leonardo Midolo[1], Sven Scholz[2], Andreas Dirk Wieck[2], Arne Ludwig[2], Anders Søndberg Sørensen[1], Peter Lodahl[1]*

Photon emission is the hallmark of light-matter interaction and the foundation of photonic quantum science, enabling advanced sources for quantum communication and computing. Although single-emitter radiation can be tailored by the photonic environment, the introduction of multiple emitters extends this picture. A fundamental challenge, however, is that the radiative dipole-dipole coupling rapidly decays with spatial separation, typically within a fraction of the optical wavelength. We realize distant dipole-dipole radiative coupling with pairs of solid-state optical quantum emitters embedded in a nanophotonic waveguide. We dynamically probe the collective response and identify both super- and subradiant emission as well as means to control the dynamics by proper excitation techniques. Our work constitutes a foundational step toward multiemitter applications for scalable quantum-information processing.

The radiative coupling of multiple optical emitters has been a long-standing challenge in quantum optics and atomic physics (*1*–*4*). It offers a route to realizing quantum gates between emitters (*5*), thereby constituting a fundamental building block for quantum-information processing (*6*). Waveguide quantum electrodynamics (QED) (*7*) has evolved as a research discipline ideally suited for overcoming the inherently weak dipole-dipole coupling. This is because the radiative coupling is extended substantially beyond the subwavelength limit encountered in unstructured media (*8*). The dipole-dipole interaction can be understood as the absorption and re-emission of virtual photons, and the waveguide extends the spatial range to be limited only by the weak leakage of the waveguide mode because of structural imperfections (*9*). This leads to the formation of collective emitter states featuring super- or subradiant decay rates (*1*, *2*, *10*), as controlled by the optical phase lag between the two emitters (Fig. 1A). Observation of collective emission requires a highly coherent light-matter interface. In particular, long-lived subradiant features may be elusive in the presence of experimental imperfections, such as dephasing (*11*).

Collective multiemitter effects have been studied in the optical (*3*, *4*, *8*, *12*–*15*) and microwave domains (*16*–*18*). In microwave QED, multiqubit interactions have been realized (*19*, *20*) and subradiant collective states coherently controlled (*21*). Realizing such functionalities in the optical domain is essential: Optical photons can be highly integrated, rapidly processed on-chip, and transmitted over extended distances (*6*, *22*), which makes photonics the backbone technology for the quantum internet (*23*). Previous accounts in the optical domain include projective preparation of single-excitation superradiant states, as revealed in two-photon correlation measurements (*24*–*27*) that do not require emitter-emitter coupling. Spectroscopic evidence for coherent coupling has been reported with closely spaced dye molecules in a bulk medium (*8*, *28*) and for vacancy centers in a cavity (*15*). Dynamics with modified collective emission would constitute direct experimental proof of coherent coupling and open new avenues toward applications. Here, we report the observation of coherent dynamics of the collective excitation of quantum dot (QD) emitters coupled through a photonic crystal waveguide (PCW) (Fig. 1A) by observing both enhanced (superradiant) and suppressed (subradiant) dynamics. In contrast to cooperative (*29*) and amplified (*30*) spontaneous emission involving multiple excitations, we concentrate on a single excitation distributed between a pair of QDs. In the former case, the resulting emission peak intensity scales as N^2, whereas in the latter, the enhancement is $\propto N$, where N is the number of emitters involved. In the experiment, the QDs are brought into mutual resonance by Zeeman tuning with a magnetic field. The coherent oscillations of the collective state are directly observed and found to be controllable by spectrally tuning the QDs and by varying the excitation conditions.

[1]Center for Hybrid Quantum Networks (Hy-Q), The Niels Bohr Institute University of Copenhagen, DK-2100 Copenhagen Ø, Denmark. [2]Lehrstuhl für Angewandte Festkörperphysik, Ruhr-Universität Bochum, Universitätsstraß e 150, D-44801 Bochum, Germany.
***Corresponding author. Email: alexey.tiranov@nbi.ku.dk (A.T.); lodahl@nbi.ku.dk (P.L.)**
†These authors contributed equally to this work.

Hamiltonian for waveguide-mediated interaction

The coupled QDs are described by the effective Hamiltonian $\mathcal{H}_{\text{eff}}$ (frame rotating at the excitation field frequency and $\hbar = 1$) (*31*, *32*)

$$\mathcal{H}_{\text{eff}} = \sum_{m,n=1}^{2}\left(J_{mn} - i\frac{\Gamma_{mn}}{2}\right)\sigma_n^+\sigma_m^- + \sum_{n=1}^{2}\left(\Delta_n - i\frac{\gamma_n^s}{2}\right)\sigma_n^+\sigma_n^- \quad (1)$$

where $J_{mn} = \frac{1}{2}\sqrt{\beta_m\beta_n\Gamma_m\Gamma_n}\sin\phi_{mn}$ and $\Gamma_{mn} = \sqrt{\beta_m\beta_n\Gamma_m\Gamma_n}\cos\phi_{mn}$ are the dispersive and dissipative coupling rates, respectively, connecting QD_m and QD_n. $\Gamma_m = \gamma_m^{wg} + \gamma_m^s$ contains the decay rate into (and out of) the waveguide γ_m^{wg} (γ_m^s), corresponding to β factors $\beta_m = \gamma_m^{wg}/\Gamma_m$. Δ_m is the detuning of QD_m with respect to the excitation field frequency, such that $\Delta_{mn} = \Delta_m - \Delta_n$ is the detuning between the two QDs, $\phi_{mn} = k|x_{mn}|$ is the phase lag resulting from the emitter separation x_{mn}, and k is the effective wave number of the PCW mode. ϕ_{mn} determines the character of the coupling between dispersive ($\Gamma_{mn} = 0$), which modifies the energy levels, and dissipative ($J_{mn} = 0$), which affects the decay dynamics. The system bears a resemblance to the case of a nanocavity, where each QD acts as an end mirror by scattering single photons into the

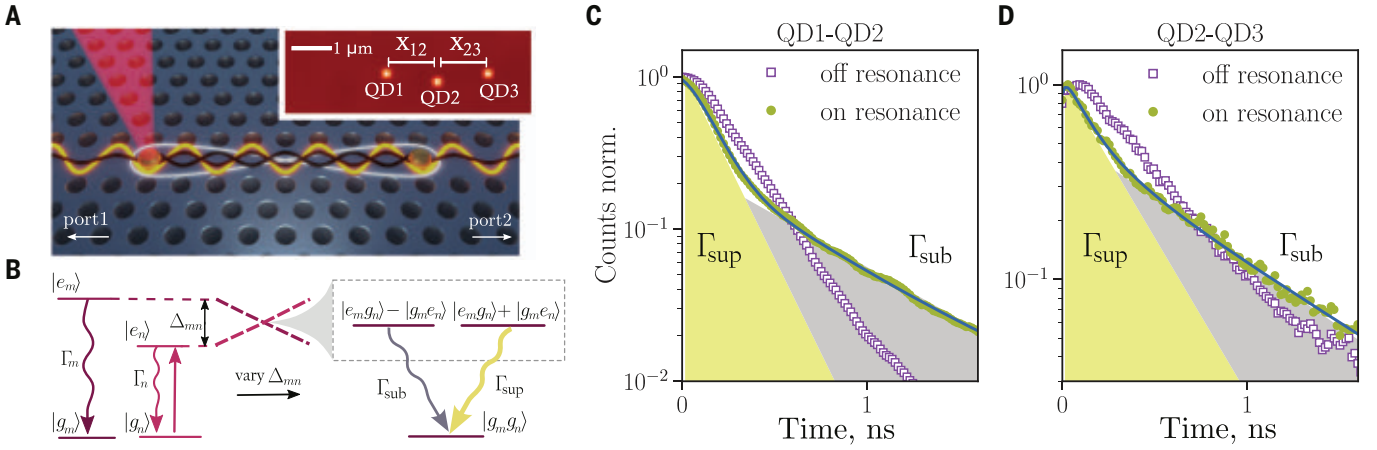


Fig. 1. Observation of the super- and subradiant emission. (**A**) Illustration of the photon-mediated coupling between two QDs in a PCW where one QD is optically pumped. Subsequently, the emission dynamics of the coupled QD system exhibit super- and subradiance originating from either constructive (bright line) or destructive (dark line) interference of the field emitted into the PCW and scattered by each of the QDs. (Inset) Fluorescence image of the QDs in the PCW, where x_{12} = 1.25(3) μm and x_{23} = 0.96(3) μm. (**B**) Energy-level diagram of two QDs with a mutual detuning of Δ_{mn}. Off resonance (left), the two QDs decay independently of each other, whereas on resonance (right), super- and subradiant dynamics occur with rates Γ_{sup} and Γ_{sub}, respectively. (**C** and **D**) Measured time-resolved emission dynamics of pair QD1-QD2 (C) and QD2-QD3 (D) both on and off resonance and after exciting QD2 and QD3 through higher-energy excited states. The measured count rates are normalized to the maximum at zero time delay, corresponding to the time of excitation of the coupled system. The solid lines are the model fits to the experimental data. Off resonance, $\Delta_{12}/2\pi$ = 6 GHz ($\Delta_{23}/2\pi$ = 5 GHz) the decay curve corresponds to the cases of individual QDs. When two QDs are brought into resonance, two decay components are observed in the data, a fast Γ_{sup} and a slow Γ_{sub} corresponding to super- and subradiant dynamics.

mode of the waveguide, and ϕ_{mn} determines the cavity resonance condition (Fig. 1A). Notably, the system is inherent in quantum character because the QDs only scatter a single photon at a time. σ_m^+ and σ_n^- are the raising and lowering operators for the optical transition of QD_m and QD_n, respectively.

By selectively pumping one of the QDs—e.g., QD_n—one excitation is launched to populate the state $|g_m e_n\rangle$. On resonance (Δ_{mn} = 0), the subsequent dynamics is best described in terms of the entangled eigenstates $|S\rangle = (|e_m g_n\rangle + |g_m e_n\rangle)/\sqrt{2}$ and $|s\rangle = (|e_m g_n\rangle - |g_m e_n\rangle)/\sqrt{2}$ with the associated decay rates, determined by the phase lag ϕ_{mn}. The case when $\phi_{mn} = N\pi$ (N integer) corresponds to a dissipative coupling between QDs leading to a superradiant $|S\rangle$ and a subradiant $|s\rangle$ state with modified decay rates (Fig. 1B). $\phi_{mn} = (N + {}^1/_2)\pi$ results in dispersive coupling, where the collective states are shifted by J_{mn} while leaving their decay rates unchanged compared with those of the uncoupled emitters. In the present experiment, we are primarily studying the regime of dissipative radiative coupling, which leads to modified emission dynamics.

Super- and subradiance with coherent evolution

We optically excite a single QD and record the collective emission dynamics from either collection port 1 or 2, studying three different QDs (QD1 to QD3; Fig. 1A, inset). Two examples of recorded emission signals are shown in Fig. 1, C and D, where we alter the detuning to compare the off- and on-resonant cases; see also the supplementary text for further experimental details (*33*). On resonance, we observe strongly modified decay dynamics resulting from coherent coupling, and both super- and subradiant features are directly visible. For QD1-QD2, we find radiative linewidths of $\Gamma_{sup}/2\pi$ = 1.36(8) GHz and $\Gamma_{sub}/2\pi$ = 0.280(2) GHz by modeling the data with a biexponential decay. The modeling of the data at short time delays is limited by the finite instrument response function of the single-photon detectors, particularly visible at the rising edge of the detected pulse. These values should be referred to the single emitter linewidths of $\Gamma_2/2\pi$ = 0.79(2) GHz and $\Gamma_1/2\pi$ = 0.85(1) GHz, respectively, as observed far off resonance, where the coupling is negligible. We derive a superradiant/subradiant enhancement factor of 1.36/0.28 = 4.8. The enhancement factor is a direct figure of merit of the collective coupling quality and is highly sensitive to experimental imperfections and decoherence [for a detailed account of the underlying physical parameters, see supplementary text, table S2 (*33*)]. The long-range nature of the dipole-dipole coupling is explicitly demonstrated. Using resonant excitation through the PCW, we image the spatial separation of the QDs and find x_{12} = 1.25(3) μm and x_{23} = 0.96(3) μm (Fig. 1A, inset), which should be compared with the wavelength of λ = 270 nm inside GaAs or the PCW lattice constant of a = 240 nm.

Next, we show the coherent evolution of collective states by precisely controlling the detuning between two emitters. We start by resonantly exciting a single QD to prepare $|eg\rangle = (|S\rangle + |s\rangle)/\sqrt{2}$—i.e., an equal superposition of the super- and subradiant states. The state subsequently evolves in time into a collective state (see the Bloch sphere graphical representation in Fig. 2, A and B). The color of the Bloch sphere surface represents the decay time of the respective state, and the collective state vector precesses as a result of coherent evolution. Assuming $\Gamma_m = \Gamma_n$, the correlated dynamics is described by the difference between the two eigenvalues of the coupled system $f_{osc} = \sqrt{\Delta_{mn}^2 + (2J_{mn} - i\Gamma_{mn})^2}$. For example, in the case of pure dissipative coupling (J_{mn} = 0), two different regimes are identified (Fig. 2, C and D). In the underdamped regime $|\Delta_{mn}| > \Gamma_{mn}$, coherent evolution prevails over dissipation, resulting in the observation of an increase in emission intensity. The dashed lines in Fig. 2, D to F, track the observed intensity maximum for different values of Δ_{mn} defined by $t = \pi/f_{osc}$. In the overdamped regime $|\Delta_{mn}| < \Gamma_{mn}$, the dissipative coupling damps the excitation to the ground state faster than the coherent oscillations on the Bloch sphere (leading to the gap between the two dashed lines in Fig. 2, D to F). $\Delta_{mn} = \Gamma_{mn}$ is the case of critical damping, where the dissipation and coherent oscillation rates are balanced. These examples quantify how precise emitter tuning provides an experimentally accessible "control knob" of coupled collective quantum states.

The rich coherent dynamics of the collectively coupled system are evident from the

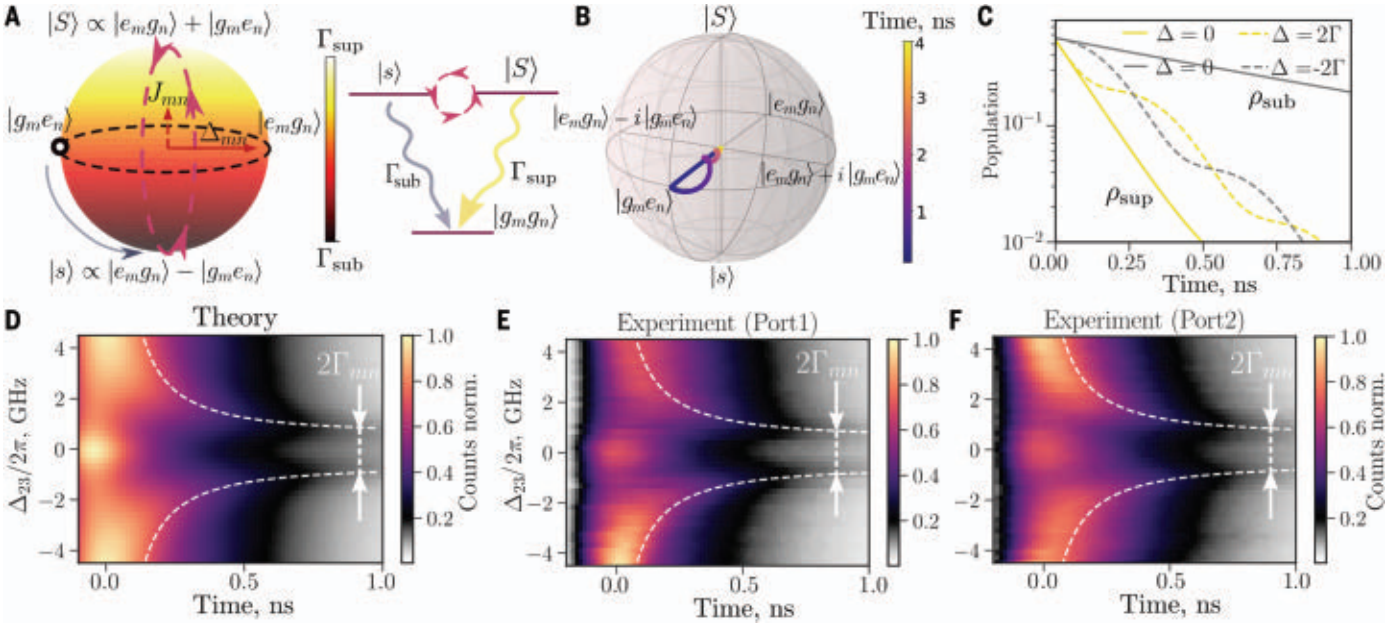


Fig. 2. Coherent dynamics of coupled QDs. (**A**) The evolution of a single collective excitation can be represented on a Bloch sphere, with the color indicating the decay rate of the respective collective state. After exciting QD_n, the state $|g_m e_n\rangle$ is populated. Two processes occur: exponential decay by spontaneous emission (black solid arrow) and coherent evolution between super- and subradiant states (red arrow) as determined by the detuning between the two QDs Δ_{mn}. (**B**) Exemplary state evolution trajectory for $\Delta_{23} = -5.5\Gamma$, where the length of the Bloch vector shows the population in the single-excitation subspace. The color bar tracks the evolution time from preparing the coupled system in the initial state until decaying to the ground state $|g_m g_n\rangle$ (origin of Bloch sphere). (**C**) Calculated population of the collective states as a function of time for two values of detuning and comparing superradiant (yellow curves) and subradiant (gray curves) contributions for an initial state $|g_m e_n\rangle$. (**D**) Plot of the corresponding emission intensity. The calculation is done with the experimentally extracted parameters of the QD2-QD3 pair (see supplementary text). (**E** and **F**) Full experimental dataset for a continuous scan of detuning Δ_{23} and comparing measurements from outcoupling port 1 (E) and port 2 (F). The dashed lines trace the maximum population after the first oscillation and are given by $t = \pi/f_{osc}$ (see the text for definitions).

experimental data (Fig. 2, E and F) that are well reproduced by theory (Fig. 2D). As for the resonant case, after a fast decay of the superradiant component, the coupled system evolves toward the subradiant state. In the detuned case, coherent evolution is observed, leading to modified dynamics because super- and subradiant components are interchanged. The oscillation between super- and subradiant components, as defined by the detuning, leads to an increase of the intensity, corresponding to the superradiant component being maximal. This is clearly visible as a bright region, as it is tracked by the dashed line in Fig. 2, D to F. By comparing the emission from two different directions (outcoupling ports 1 and 2), we observe a similar behavior (Fig. 2, E and F), which is consistent with a predominantly dissipative coupling $\phi_{mn} \simeq N\pi$. The experiment was repeated on three pairs of QDs; the additional data and detailed modeling can be found in the supplementary text (*33*). For all three datasets, we observe predominantly dissipative coupling, which likely results from the fact that QD candidates, featuring efficient coupling to the PCW, were preselected in the experiment on the basis of resonant transmission measurements through the waveguide [see supplementary text (*33*)]. This condition results in the selection of QDs close to the waveguide center and, therefore, a phase lag between QDs primarily determined by the periodicity of the photonic crystal lattice. Because the PCW coupling is large and QDs are spectrally close to the band edge of the waveguide mode, this leads to $k/a \simeq \pi$ (*9*), whereby the phase lag between neighboring PCW unit cells is $\simeq \pi$ [see the supplementary text for further details (*33*)].

The theoretical model (Fig. 2D) reproduces the experimental data and fully captures the complex coherent quantum dynamics observed experimentally (Fig. 2, E and F). From the analysis, we extract for pair QD2-QD3, $\Gamma_{23}/2\pi$ = 0.61 GHz and $J_{23}/2\pi$ = 0.03 GHz. The dissipative coupling rate is comparable to the intrinsic linewidth of the respective QDs, whereas the dispersive part is almost vanishing. In the case of negligible dispersive coupling, the superradiant Γ_{sup} (and subradiant Γ_{sub}) decay rate can be approximated by $\Gamma_{sup} \approx \Gamma + \Gamma_{ij} - \sigma_{sd}^2/2\beta\Gamma$ ($\Gamma_{sub} \approx \Gamma - \Gamma_{ij} + \sigma_{sd}^2/2\beta\Gamma$), where σ_{sd} is the spectral diffusion width. The predicted value of $\Gamma_{sup}/2\pi = 1.25$GHz ($\Gamma_{sub}/2\pi = 0.27$GHz) for the QD2-QD3 pair is in agreement with the experimentally measured $\Gamma_{sup}/2\pi = 1.33$GHz($\Gamma_{sub}/2\pi = 0.22$GHz). For this QD pair, we obtain β_2 = 0.88 and β_3 = 0.83 from the resonant transmission data [see the supplementary materials (*33*) for further details].

As opposed to the case of superconducting qubits (*16*), changing the detuning between the QDs has a negligible effect on the phase lag, whereby the coupling remains dissipative. This is a consequence of the fact that the detuning is vanishingly small compared with the optical frequency. This distinction may be advantageous in applications of radiative collective coupling because it allows the detuning to be exploited as a control parameter without changing the coupling, which is explicitly demonstrated in the present work. Notably, even when the emitter-emitter system is initialized in $|eg\rangle$ or $|ge\rangle$, coherent oscillations are still observed. This is enabled by the fast decay of the superradiant component after initialization, leading to the population of the slower-decaying subradiant state that coherently evolves on the Bloch sphere (Fig. 2B).

Control of collective excitations

The deterministic preparation of collective states is essential to pave the way for their applications in quantum-information processing. To this end, we coherently excite both QDs to control the initial collective state on

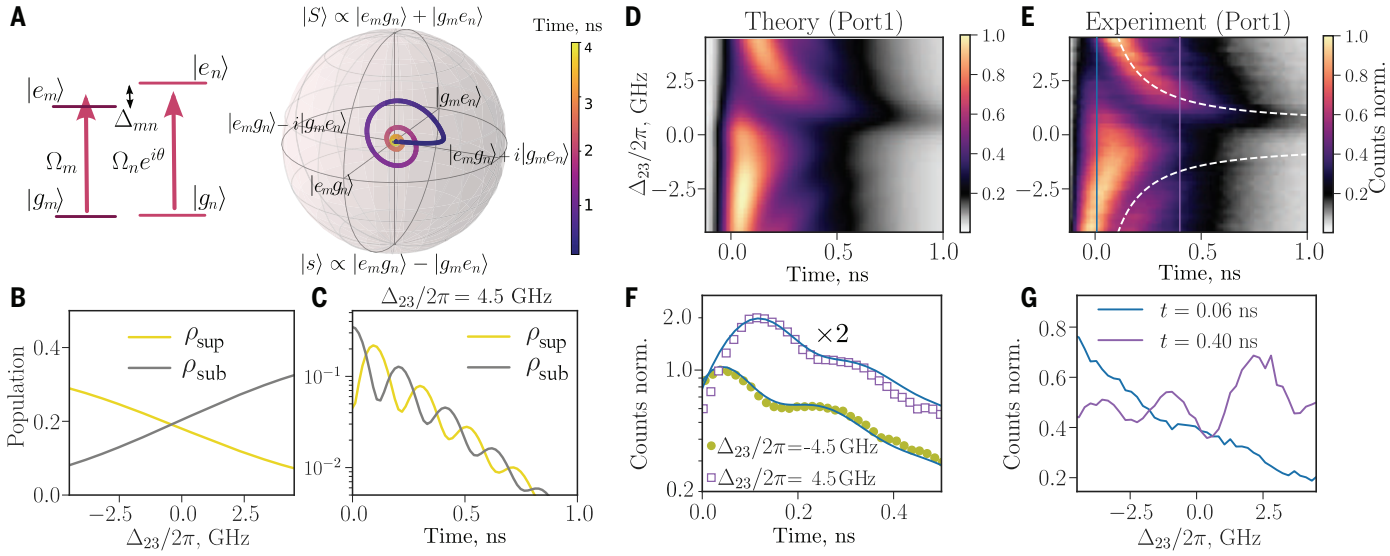


Fig. 3. Controlled preparation of the collective state. (**A**) Simultaneous driving of both QDs from the QD2-QD3 pair with control of the phase θ allows us to initialize the system in a collective state. In the case of $\theta \approx -\pi/2$, the initial state in the single-excitation subspace is close to $|ge\rangle + i|eg\rangle$. The detuning between the QDs results in the evolution of the collective state toward $|S\rangle$ when $\Delta_{23} < 0$ and toward $|s\rangle$ for $\Delta_{23} > 0$. An example of the state trajectory for $\Delta_{23} < 0$ is depicted on the Bloch sphere. (**B**) The population of the super- ρ_{sup} and subradiant ρ_{sub} components during the excitation as a function of the detuning Δ_{23} for $\theta \approx -\pi/2$. The sign of the detuning determines whether the super- or subradiant states are preferentially populated. (**C**) The two populations oscillate out of phase as a function of time after excitation. (**D** and **E**) Calculated (D) and measured (E) intensity decay dynamics as a function of detuning, collected from the outcoupling port 1. The asymmetric dependence on detuning is well reproduced by the theory using parameters extracted from the experiment with $\theta \approx -\pi/2$. (**F**) Examples of emission intensity decay curves for two values of detuning explicitly displaying the coherent oscillations being out-of-phase with each other, as determined by the sign of the detuning. (**G**) Normalized counts as a function of detuning and for two representative times quantifying the asymmetry.

the Bloch sphere (Fig. 3A). In the magnetic Zeeman field, the coupled QD transitions are orthogonally (circularly) polarized (*34*); yet, they are efficiently coupled by the optical mode of the PCW. The phase between the two driving fields Ω_2 and $\Omega_3 e^{i\theta}$ is adjusted through the polarization of the excitation laser. Using a single laser, we implement $\theta \approx -\pi/2$, which prepares an initial state with the single excitation close to $|eg\rangle + i|ge\rangle$ for $\Delta_{23} = 0$. With detuning, the state either evolves toward the super- (for $\Delta_{23} < 0$) or the subradiant (for $\Delta_{23} > 0$) state. This results in a marked difference in the radiation at a short time, followed by out-of-phase coherent oscillations between the two components (Fig. 3, B and C).

The experimental demonstration of this behavior (Fig. 3E) is accurately described by the theory (Fig. 3D). We observe a pronounced asymmetry around zero detuning, where for positive detunings, the emission dynamics is effectively delayed. This stems from the selective population of the subradiant state, resulting in a lower emission at early times. Subsequently, the emission intensity increases as the coherent evolution increases the population of the superradiant state. The reverse behavior is found for $\Delta_{23} < 0$. This is a result of the out-of-phase oscillation of the population of the super- and subradiant components (Fig. 3, F and G). As opposed to the case where a single QD was excited (Fig. 2), we observed multiple coherent oscillations. This is because of the state $|eg\rangle + i|ge\rangle$ starting to coherently evolve on the Bloch sphere directly after excitation. This is in contrast to the case where $|eg\rangle$ is prepared, and the coherent evolution sets in once the state has partially decayed to $|s\rangle$. We track the maximum emission intensity (dashed line) for $\Delta_{23} > 0$ (and the minimum for $\Delta_{23} < 0$) in the plots of Fig. 3E. It is consequently found that the collective light emission intensity can be controlled through the QD detuning. By simultaneous driving, we populate the doubly excited state $|ee\rangle$ component, which reaches 0.07 for $\approx \pi/3$ excitation pulses that are used. We note that the doubly excited state $|ee\rangle$, however, does not contribute to the measured asymmetry between positive and negative detunings.

Conclusions

Our observation of super- and subradiant emission dynamics using pairs of QDs embedded in PCWs and separated by a distance much larger than the wavelength is facilitated by the PCW offering broadband spectral operation and long-range photon-mediated coherent interaction between QDs. Our work can constitute a foundational step toward multiemitter applications of technological importance—e.g., for realizing quantum transduction between microwave qubits and the optical domain (*35*) or for quantum memories with exponential improvement in photon storage fidelity (*31*). The ability to manipulate the super- and subradiant state dynamics by controlling the detuning and pumping conditions will lead to a whole range of opportunities when implementing a coherent spin inside the QD (*34*, *36*). For example, advanced photonic cluster states may be generated deterministically (*37*), providing a universal resource for measurement-based photonic quantum computing. To this end, waveguide-mediated dissipative coupling can be exploited to realize spin-spin entanglement between distant quantum emitters (*38*). Another direction will exploit photon scattering from coupled QDs to realize efficient Bell-state measurements (*39*) or photon-photon quantum gates (*40*). Taking a broader perspective, the ability to deterministically couple multiple quantum emitters opens a previously unexplored arena of studying nonequilibrium quantum many-body physics of strongly correlated light and matter, which could be used in quantum simulations of strongly correlated condensed-matter systems (*41*).

REFERENCES AND NOTES

1. R. H. Dicke, *Phys. Rev.* **93**, 99–110 (1954).
2. M. Gross, S. Haroche, *Phys. Rep.* **93**, 301–396 (1982).
3. R. G. DeVoe, R. G. Brewer, *Phys. Rev. Lett.* **76**, 2049–2052 (1996).
4. J. Eschner, C. Raab, F. Schmidt-Kaler, R. Blatt, *Nature* **413**, 495–498 (2001).

5. A. Beige, D. Braun, P. L. Knight, *New J. Phys.* **2**, 22 (2000).
6. J. L. O'Brien, A. Furusawa, J. Vučković, *Nat. Photonics* **3**, 687–695 (2009).
7. D. E. Chang, J. S. Douglas, A. González-Tudela, C.-L. Hung, H. J. Kimble, *Rev. Mod. Phys.* **90**, 031002 (2018).
8. J.-B. Trebbia, Q. Deplano, P. Tamarat, B. Lounis, *Nat. Commun.* **13**, 2962 (2022).
9. P. Lodahl, S. Mahmoodian, S. Stobbe, *Rev. Mod. Phys.* **87**, 347–400 (2015).
10. M. O. Scully, *Phys. Rev. Lett.* **115**, 243602 (2015).
11. X.-L. Chu, V. Angelopoulou, P. Lodahl, N. Rotenberg, *Phys. Rev. A* **106**, 053702 (2022).
12. A. Goban *et al.*, *Phys. Rev. Lett.* **115**, 063601 (2015).
13. A. Sipahigil *et al.*, *Science* **354**, 847–850 (2016).
14. P. Solano, P. Barberis-Blostein, F. K. Fatemi, L. A. Orozco, S. L. Rolston, *Nat. Commun.* **8**, 1857 (2017).
15. R. E. Evans *et al.*, *Science* **362**, 662–665 (2018).
16. A. F. van Loo *et al.*, *Science* **342**, 1494–1496 (2013).
17. J. A. Mlynek, A. A. Abdumalikov, C. Eichler, A. Wallraff, *Nat. Commun.* **5**, 5186 (2014).
18. M. Mirhosseini *et al.*, *Nature* **569**, 692–697 (2019).
19. J. D. Brehm *et al.*, *npj Quantum Mater.* **6**, 10 (2021).
20. E. Kim *et al.*, *Phys. Rev. X* **11**, 011015 (2021).
21. M. Zanner *et al.*, *Nat. Phys.* **18**, 538–543 (2022).
22. R. Uppu, L. Midolo, X. Zhou, J. Carolan, P. Lodahl, *Nat. Nanotechnol.* **16**, 1308–1317 (2021).
23. H. J. Kimble, *Nature* **453**, 1023–1030 (2008).
24. J.-H. Kim, S. Aghaeimeibodi, C. J. K. Richardson, R. P. Leavitt, E. Waks, *Nano Lett.* **18**, 4734–4740 (2018).
25. J. Q. Grim *et al.*, *Nat. Mater.* **18**, 963–969 (2019).
26. Z. X. Koong *et al.*, *Sci. Adv.* **8**, eabm8171 (2022).
27. J. Q. Grim *et al.*, *Phys. Rev. B* **106**, L081403 (2022).
28. C. Hettich *et al.*, *Science* **298**, 385–389 (2002).
29. R. H. Lehmberg, *Phys. Rev. A* **2**, 889–896 (1970).
30. L. Allen, G. I. Peters, *Phys. Rev. A* **8**, 2031–2047 (1973).
31. A. Asenjo-Garcia, M. Moreno-Cardoner, A. Albrecht, H. J. Kimble, D. E. Chang, *Phys. Rev. X* **7**, 031024 (2017).
32. A. Albrecht *et al.*, *New J. Phys.* **21**, 025003 (2019).
33. See the supplementary materials containing additional details on the theoretical model, experimental measurements, and simulations.
34. R. J. Warburton, *Nat. Mater.* **12**, 483–493 (2013).
35. V. E. Elfving, S. Das, A. S. Sørensen, *Phys. Rev. A* **100**, 053843 (2019).
36. M. H. Appel *et al.*, *Phys. Rev. Lett.* **128**, 233602 (2022).
37. S. E. Economou, N. Lindner, T. Rudolph, *Phys. Rev. Lett.* **105**, 093601 (2010).
38. M. Gullans *et al.*, *Phys. Rev. Lett.* **109**, 235309 (2012).
39. J. Borregaard, A. S. Sørensen, P. Lodahl, *Adv. Quantum Technol.* **2**, 1800091 (2019).
40. B. Schrinski, A. S. Sørensen, Polariton interaction in one-dimensional arrays of atoms coupled to waveguides. arXiv:2110.15878 [quant-ph] (2021).
41. C. Noh, D. G. Angelakis, *Rep. Prog. Phys.* **80**, 016401 (2017).
42. A. Tiranov *et al.*, Data and analysis files for the publication Collective super- and subradiant dynamics between distant optical quantum emitters, Electronic Research Data Archive (Univ. Copenhagen, 2022); https://doi.org/10.17894/ucph.eea76cf8-f7fd-408c-b191-b60fbd7cc180.

ACKNOWLEDGMENTS

The authors thank X.-L. Chu for fruitful discussions at the beginning of the project. **Funding:** We gratefully acknowledge financial support from Danmarks Grundforskningsfond (DNRF 139, Hy-Q Center for Hybrid Quantum Networks). B.S. acknowledges financial support from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), grant no. 449674892. O.A.D.S. acknowledges funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 801199. S.S., A.L., and A.D.W. acknowledge financial support of the German-French University DFH/UFA within the CDFA-05-06 as well as BMBF QR.X project 16 KISQ 009. **Author contributions:** A.T., V.A., and C.J.v.D. carried out the measurements. Y.W. and L.M. designed and fabricated the sample. O.A.D.S., B.S., and A.S.S. developed the theoretical model. A.T., V.A., and C.J.v.D. analyzed the data and prepared the figures. S.S., A.D.W., and A.L. carried out the growth and design of the wafer. A.T., V.A., C.J.v.D., and P.L. wrote the manuscript with input from all the authors. P.L. supervised the project. **Competing interests:** P.L. is founder of the company Sparrow Quantum, which commercializes single-photon sources. The authors declare no other competing interests. **Data and materials availability:** The data presented in the main text and supplementary materials for this publication are openly available online (*42*).

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade9324
Supplementary Text
Figs. S1 to S18
Tables S1 to S3
References (*43*–*49*)

Submitted 21 September 2022; accepted 22 December 2022
10.1126/science.ade9324

EVOLUTIONARY ECOLOGY

Rapid plant trait evolution can alter coastal wetland resilience to sea level rise

M. L. Vahsen[1]*, M. J. Blum[2], J. P. Megonigal[3], S. J. Emrich[2,4], J. R. Holmquist[3], B. Stiller[1], K. E. O. Todd-Brown[5], J. S. McLachlan[1]*

Rapid evolution remains a largely unrecognized factor in models that forecast the fate of ecosystems under scenarios of global change. In this work, we quantified the roles of heritable variation in plant traits and of trait evolution in explaining variability in forecasts of the state of coastal wetland ecosystems. A common garden study of genotypes of the dominant sedge *Schoenoplectus americanus*, "resurrected" from time-stratified seed banks, revealed that heritable variation and evolution explained key ecosystem attributes such as the allocation and distribution of belowground biomass. Incorporating heritable trait variation and evolution into an ecosystem model altered predictions of carbon accumulation and soil surface accretion (a determinant of marsh resilience to sea level rise), demonstrating the importance of accounting for evolutionary processes when forecasting ecosystem dynamics.

Organismal traits have long been understood to drive ecosystem functions such as elemental cycling (*1*, *2*). There is mounting evidence that heritable trait variation within species can mediate ecosystem processes at a magnitude comparable with that of trait variation between species (Fig. 1B) (*3*, *4*), and that traits can evolve at a fast enough pace to generate feedbacks that alter ecosystem dynamics on the timescale of current anthropogenic environmental change (Fig. 1C) (*5*). Together, this suggests that evolutionary processes play a larger role in the regulation of ecosystem function than previously imagined (*6*–*8*) (Fig. 1). Despite growing appreciation for this possibility, efforts to predict ecosystem function that account for genetic variation and evolutionary processes remain limited in number and scope (*4*, *9*–*12*), in part because empirical studies are still needed to explicitly demonstrate whether heritable variation and rapid evolution are important drivers of ecosystem change. Studying the heritable trait variation of organisms is a necessary step toward understanding whether organismal evolution can influence ecosystem dynamics (*6*). Examining heritable trait variation over historical time might further reveal how organismal evolution elicits substantial ecosystem-level change (*13*, *14*).

In coastal marshes, dominant plants act as ecosystem engineers by contributing to soil surface accretion, a process that has allowed marshes to keep pace with sea level rise for millennia and is critical to their resilience (*15*). Further, the combination of high plant productivity and low decomposition rates from anoxic conditions in coastal marsh soils results in disproportionately high carbon accumulation rates per area relative to the soils of other ecosystems (*16*, *17*). Models and empirical syntheses have demonstrated how traits and the growth of dominant marsh plants contribute to these and other ecosystem processes (*18*, *19*). Because coastal marshes typically have low plant species diversity, intraspecific trait variation may play an important role in ecosystem processes (*12*, *20*) (Fig. 1, B and C).

Belowground plant traits exert a strong influence on marsh ecosystem processes. For example, marsh accretion responds strongly to annual root turnover, which expands marsh soils, and plant-mediated decomposition, which reduces soil volume (*21*, *22*). Belowground structures are consequently major

[1]Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, USA. [2]Department of Ecology & Evolutionary Biology, University of Tennessee, Knoxville, TN, USA. [3]Smithsonian Environmental Research Center, Edgewater, MD, USA. [4]Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN, USA. [5]Department of Environmental Engineering Sciences, University of Florida, Gainesville, FL, USA.
***Corresponding author. Email: mvahsen@nd.edu (M.L.V.); jmclachl@nd.edu (J.S.M.)**

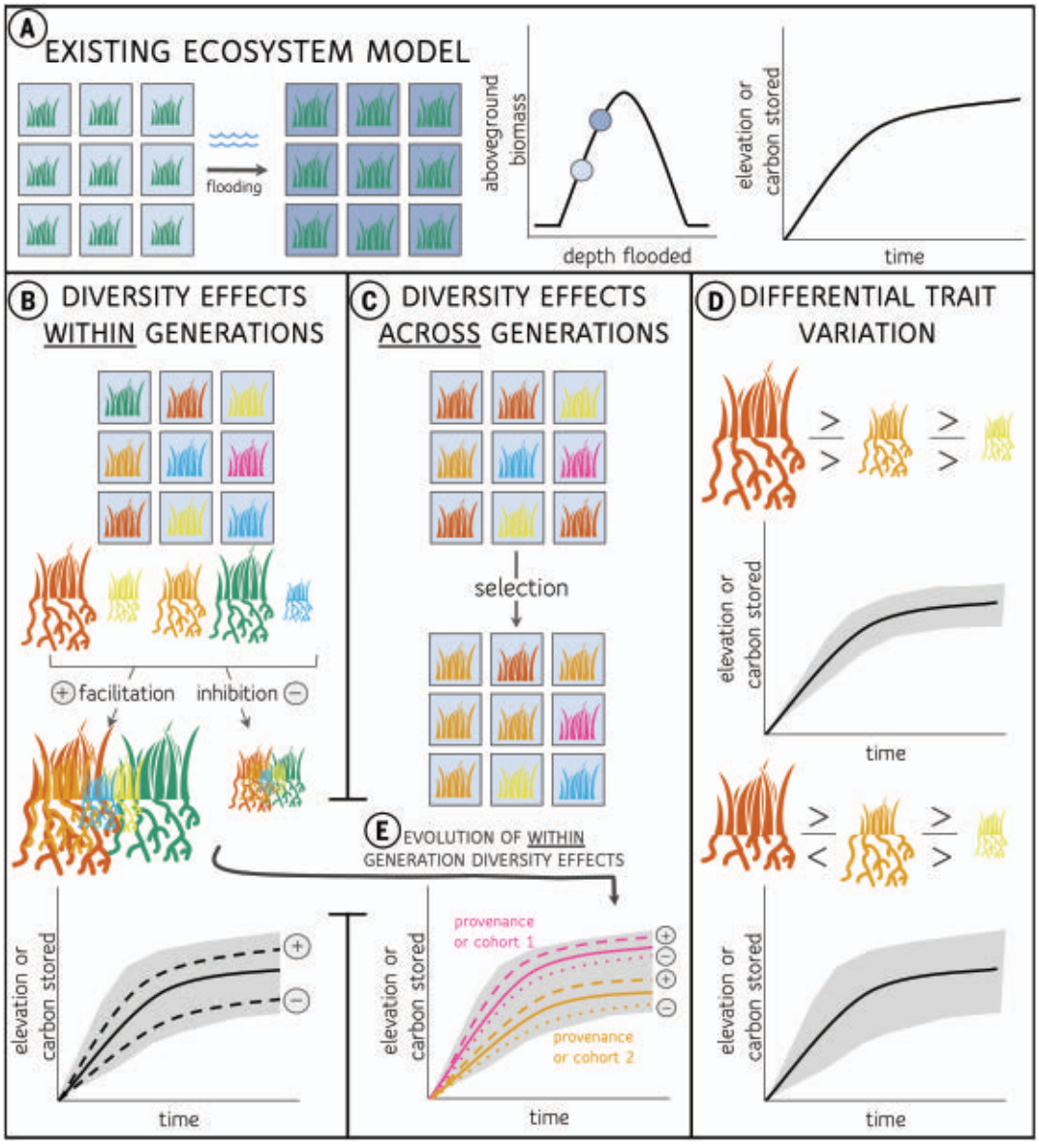


Fig. 1. Potential consequences of heritable variation in *S. americanus* on marsh accretion and carbon sequestration. (**A**) Current marsh accretion models do not account for heritable variation in traits and only account for plastic responses of aboveground biomass to flooding. Increased flooding from light to dark blue environments increases aboveground biomass. (**B**) Heritable variation may have nonadditive, within-generation consequences. For example, interactions between genotypes can lead to facilitation (+) or inhibition (−) shifting the mean trait values of polycultures and thus shifting the mean prediction for ecosystem processes. (**C**) Selection may shift plant trait means, inducing evolutionary change. (**D**) Variation in belowground traits may not scale with variation in aboveground traits; thus capturing belowground trait variation is important for accurately predicting variation in ecosystem processes. (**E**) Within-generation diversity effects [from (B)] can evolve [from (C)]. In (B) to (E), different colored plants represent different genotypes of the same species.

contributors to carbon pools, and belowground productivity is tightly linked to carbon accumulation. Empirical estimates of belowground trait variation, heritable or otherwise, are sparse (*23*, *24*), resulting in the common simplifying assumption that belowground traits vary following a fixed proportion to aboveground traits (Fig. 1D). This is a potentially unrealistic assumption—especially considering work that suggests that root-to-shoot ratios can rapidly evolve and exhibit substantial plasticity in response to stress (*25*, *26*)—that can bias predictions (*27*, *28*).

In this work, we paired a common garden experiment with an ecosystem model (*29*) to quantify the role of heritable variation in plant traits and of trait evolution in explaining variability in forecasts of carbon accumulation and soil surface accretion (Fig. 1). We characterized heritable trait variation using 16 genotypes of *Schoenoplectus americanus*—a dominant sedge in North American coastal marshes and the subject of extensive global change research related to coastal wetlands—and focused on belowground traits that are known to influence carbon sequestration and accretion. We characterized trait variation and evolution (Fig. 1, B and C) by applying a resurrection ecology approach (*14*, *30*–*32*) in which we "resurrected" genotypes from time-stratified seed banks from four nearby marshes in the Chesapeake Bay (figs. S1 to S3 and table S1). For genotypes from two of the four marshes, we assessed the role of genotype provenance (marsh of origin; Corn Island or Sellman Creek), and age cohort [ancestral (1931 to 1973) or descendant (1994 to 2016)] in driving trait variation. To assess potential nonadditive interactions that can be important when scaling up from genotype to ecosystem (Fig. 1B), we compared traits of the 16 genotypes grown in monoculture (four propagules of one genotype; n = 3 monocultures per genotype, totaling 48 monocultures) with those grown in polyculture (one propagule each of four genotypes; n = 48 total polycultures). We quantified the potential impact of eco-evolutionary dynamics on ecosystem processes using estimates of heritable variation and evolution from the common garden experiment to parameterize a marsh ecosystem model. Together, these approaches provide a framework for integrating data from common garden experiments typical in the field of evolutionary biology to predictions at the ecosystem scale appropriate for forecasting ecosystem responses to global change.

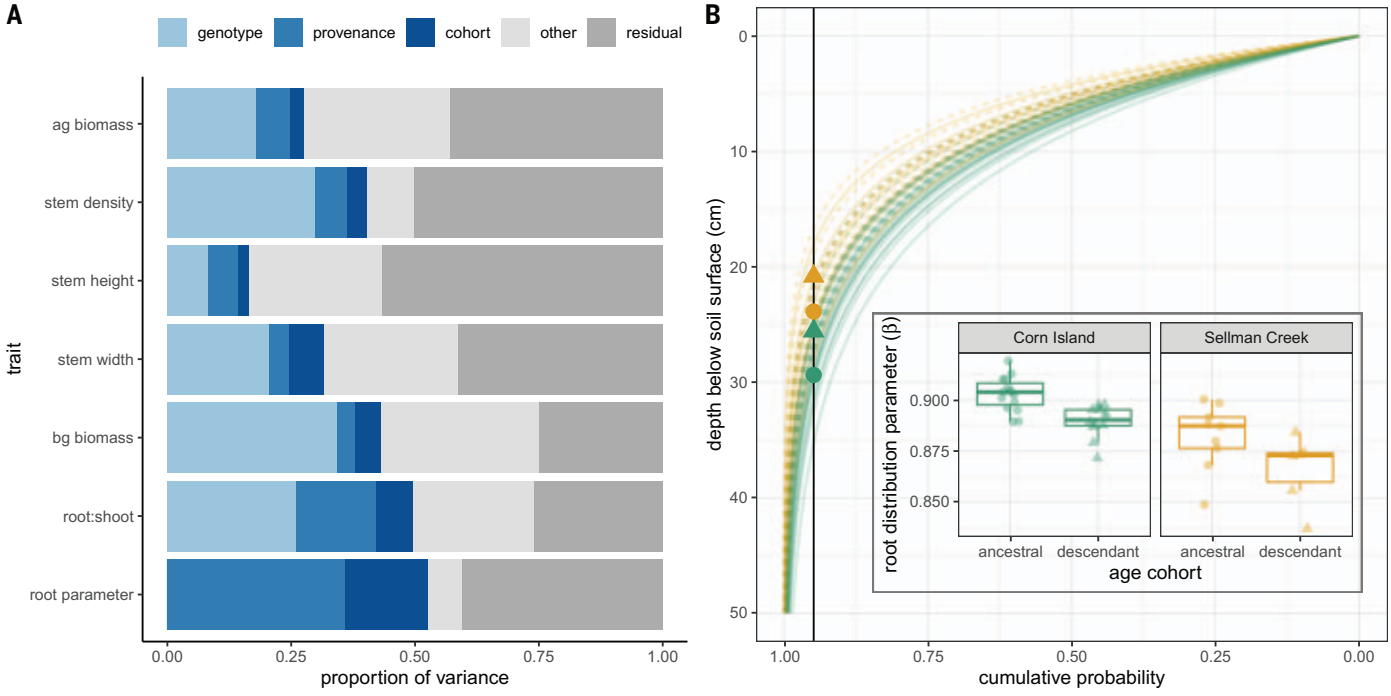


Fig. 2. Provenance and age cohort explain considerable variation in traits, particularly for root depth distribution. (**A**) Using (generalized) linear mixed models with provenance and age cohort as fixed effects and genotype as a random effect, for each trait we decomposed observed trait variation into five categories: genotype, provenance, cohort, other (covariates in the model that accounted for exogenous variation because of experimental setup: initial propagule weight, variation in flooding because of peat level, and spatial blocking), and residual. Labels "ag biomass" and "bg biomass" represent aboveground and belowground biomass, respectively. (**B**) Differences in belowground biomass distribution with depth according to provenance and age cohort. Differences in the parameter β (root distribution parameter) are shown in the inset and were applied to the equation $1 - \beta^{depth}$ to predict the cumulative proportion of belowground biomass with depth shown in the main figure (*34, 45*). The vertical line at 95% cumulative probability indicates the depth at which 95% of belowground biomass is contained, which is a parameter in the Cohort Marsh Equilibrium Model (CMEM) (*29*).

Characterizing heritable variation in *S. americanus* traits

S. americanus exhibited considerable heritable variation in all traits measured in the common garden experiment (fig. S4). Comparably more heritable variation was observed in belowground traits—such as the magnitude and distribution of belowground biomass (fig. S4, B, D, and F)—than in aboveground traits (fig. S4, A, C, E, and G). For example, heritable variation explained on average 49.5% of variation in the shape of the root depth distribution [intraclass correlation coefficient (ICC), 49.5; 95% confidence interval (CI) (15.7, 77.3)] and 69.1% of variation in root-to-shoot ratio [ICC, 69.1; (43.7, 87.7)]. These findings demonstrate the importance of explicitly characterizing variation in belowground traits, as their variation does not align with that of traits aboveground.

Heritable variation in *S. americanus* traits was structured by evolution captured across space (across the two provenances for which there were multiple genotypes; ~2 km of distance) and time (~50 years between ancestral and descendant cohorts) (Fig. 2). Differences in heritable trait variation were reflected in patterns of genetic variation elucidated with single-nucleotide polymorphism (SNP) genotyping (fig. S3 and table S3). Heritable variation attributable to genotype, provenance, and age cohort explained roughly 15 to 50% of observed variation across all traits (Fig. 2A, blue shading, and fig. S5) and, for most traits, exceeded the variation explained by experimental covariates (Fig. 2A, light gray shading; initial propagule weight, variation in flooding due to peat levels, and spatial blocking). Differences in the shape of root depth distributions were strongly consistent within age cohort and provenance (Fig. 2, A and B) [regression coefficient of age (β_{age}) = −0.015 (−0.022, −0.007); regression coefficient of provenance (β_{prov}) = −0.023 (−0.031, −0.015)]. Rooting depth became shallower over time within both provenances, with more root biomass proportionally allocated near the marsh surface in descendant genotypes (Fig. 2B). Root-to-shoot ratios exhibited strong signatures of provenance, with genotypes from Corn Island having root-to-shoot ratios that were 17.2% (1.9%, 34.4%) higher than those of genotypes from Sellman Creek (fig. S6). Comparisons of ancestral and descendant genotypes also revealed that root-to-shoot ratios have declined by 8.3% (−5.2%, 19.6%) since the mid-20th century (fig. S6), indicating that over time, plants have allocated fewer resources toward belowground biomass. We hypothesize that belowground traits may have evolved in response to anthropogenic nitrogen loading, which has increased throughout the Chesapeake Bay over the most recent century (*33*). Shifts in root depth distribution in coastal marsh vegetation have previously been posited to represent differences in how plants access nutrients belowground (*34*). Excess nitrogen may have alleviated nutrient limitation, reducing the need for plants to invest in traits that improve access to belowground resources (*35*).

Aboveground traits also evolved, but less so than belowground traits, as evidenced by smaller effect sizes of age cohort and provenance. For example, on average, stems became thinner over time, with stem widths declining 5.6% (−3.3%, 14.4%) between ancestral and descendant cohorts (fig. S7). This pattern mirrors changes in stem morphology exhibited by *S. americanus* subjected to 30 years of elevated CO_2 exposure—a change that can affect the ability of marshes to withstand storm

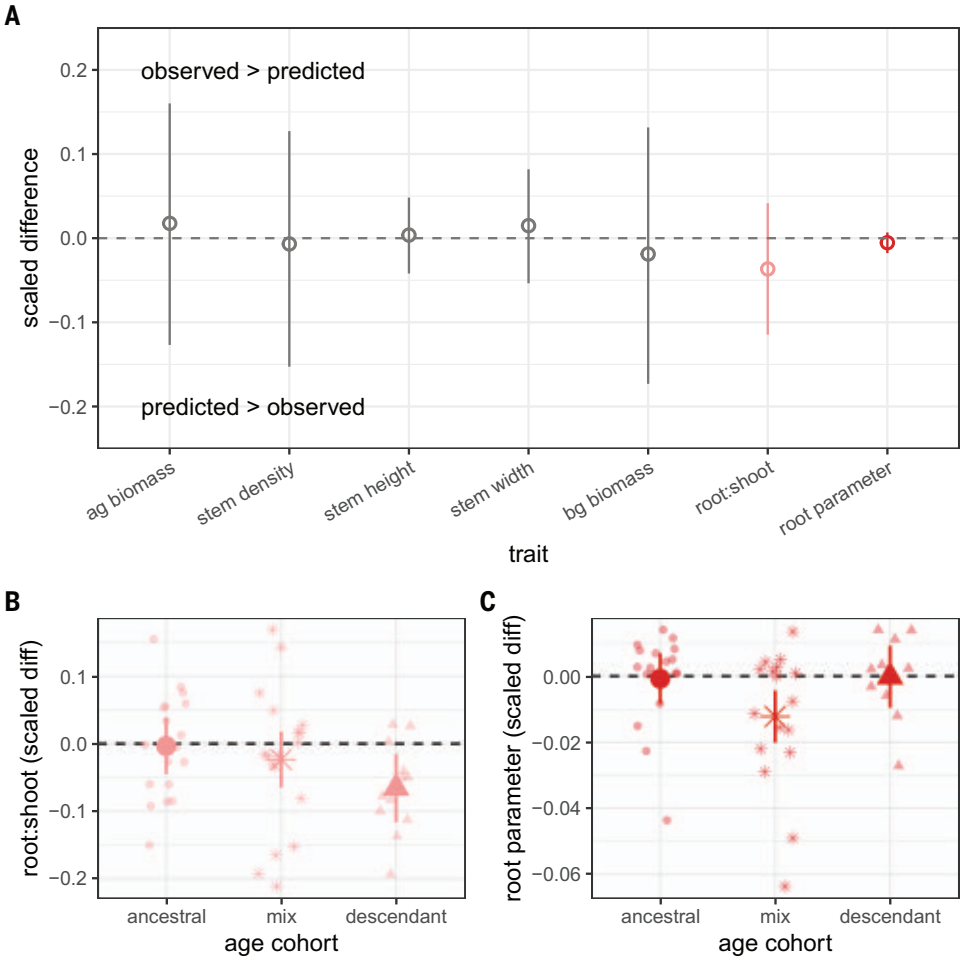


Fig. 3. Genotypes do not exhibit strong nonadditive interactions in polyculture overall, but there is evidence that within-species interactions have evolved since the mid-20th century. (**A** to **C**) The "scaled difference" on the y axes indicates the difference between a trait exhibited by genotypes grown in polyculture versus what would be expected of genotypes grown in monoculture (scaled by the mean value of the trait for easier comparison across traits). (A) Scaled differences across traits overall indicate no significant nonadditive interactions. Points indicate the mean difference, and bars indicate 95% confidence intervals. The strength of within-species interactions for (B) root-to-shoot ratio and (C) root depth distribution varies systematically based on the composition of the polyculture. For (B) and (C), points with error bars represent marginal means with 95% confidence intervals.

surges (*36*). Like belowground traits, on average, more variation in aboveground traits was attributable to provenance than age cohort. For example, mean stem heights differed according to provenance, with plants from Sellman Creek being 3.0% (−1.5%, 7.5%) taller than those from Corn Island (fig. S8), whereas only a 0.3% (−3.8%, 4.7%) difference in stem height was found between ancestral and descendant plants.

Assessing the strength of nonadditive interactions between genotypes

Given that there are high levels of standing genetic diversity within populations of *S. americanus*, even at fine spatial scales of a few meters (*30*), it is possible that interactions among genotypes result in nonadditive effects, in which trait values for a mixture of genotypes are not equal to the sum of trait values for individual genotypes (*37*–*39*) (Fig. 1B). Consequently, characterizing the direction and strength of nonadditive effects can be important for scaling trait variation from genotype to ecosystem. Mechanisms that give rise to nonadditivity can include facilitation (a positive nonadditive interaction), inhibition (a negative nonadditive interaction), and selection effects (a positive or negative nonadditive interaction) (*37*). Overall, comparisons of *S. americanus* genotypes grown in monoculture with those grown in polyculture did not reveal strong evidence of nonadditive interactions (Fig. 3A). However, for two belowground traits—root-to-shoot ratio and root depth distribution—variation in the strength of the nonadditive interactions depended on whether the polycultures were composed of ancestral genotypes, modern genotypes, or a mix of both (Fig. 3, B and C and fig S9). Root-to-shoot ratios were substantially lower than additive expectations for polycultures composed of descendant genotypes but not for those composed of ancestral genotypes [$\beta_{anc\ vs\ desc}$ = −0.06 (−0.13, 0.00)] (Fig. 3B), suggesting that the strength of within-species interactions can rapidly evolve. Root depth distributions in mixed polycultures were shallower than additive expectations, but those composed of only ancestral or descendant genotypes were similar to additive expectations [$\beta_{mix\ vs\ anc}$ = 0.01 (0.00, 0.02); $\beta_{mix\ vs\ desc}$ = 0.01 (0.00, 0.02)] (Fig. 3C). Mixing genotypes from different age cohorts in our experiment may have increased the functional diversity of rooting behavior, allowing for better resource partitioning without the need for deeper rooting (Fig. 3C).

Scaling heritable variation in plant traits to ecosystem outcomes

Observed heritable trait variation and rapid evolution drove downstream effects on soil surface accretion and carbon accumulation (Fig. 4). We ran simulations of a marsh accretion model (*29*) based on conditions at the Global Change Research Wetland (Edgewater, MD, USA), which are typical of large areas of the Chesapeake Bay. We accounted for heritable variation in peak aboveground biomass, root-to-shoot ratio, and depth of the 95% cumulative root distribution (while accounting for between-trait covariances) (figs. S10 and S11 and table S2). We projected that heritable trait variation could result in differences in marsh elevation gain of up to 5 cm by the year 2100 [mean elevation = 34.2 cm NAVD88 (North American Vertical Datum, 1988), (32.1, 37.1)]; (Fig. 4A), which is approximately one-third of the elevation differential between

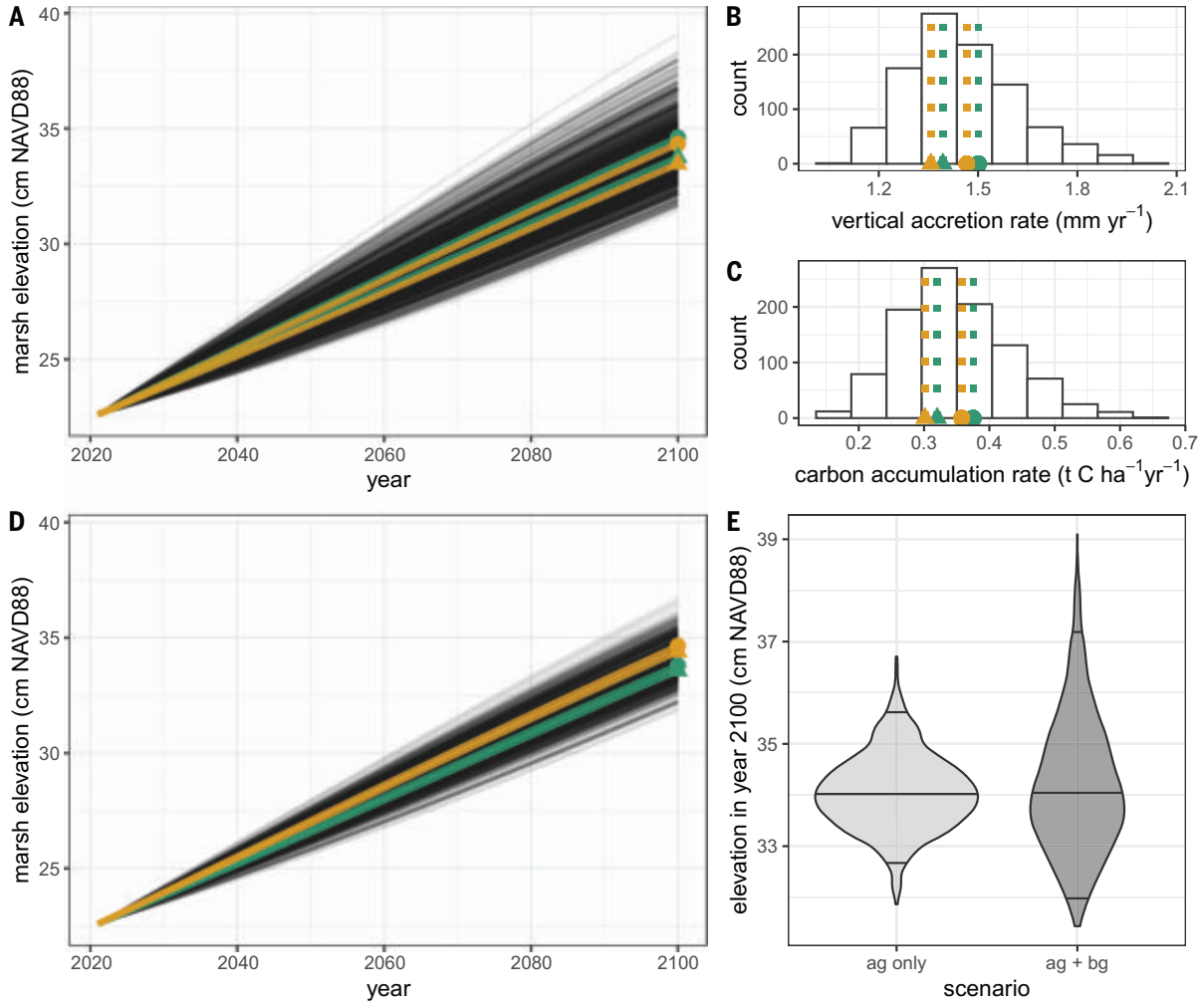


Fig. 4. Accounting for heritable trait variation and evolution alters forecasts of marsh ecosystem structure and function. (**A**) Using CMEM (*29*), we simulated marsh elevation gain to the year 2100. Light gray lines indicate model simulations (n = 1000) that account for variation in aboveground biomass, root-to-shoot ratio, and 95% cumulative root distribution depth due to genotype. Green and gold lines indicate mean predictions for genotypes from Corn Island and Sellman Creek, respectively, and the shapes at year 2100 indicate age cohorts (circle, ancestral; triangle, descendant). (**B**) Average vertical accretion rate explained by variation in traits due to heritable variation (histogram) and average provenance and age cohort trait values (points). (**C**) Average carbon accumulation rate explained by variation in traits due to heritable variation (histogram) and average provenance and age cohort trait values (points). (**D**) Simulations of CMEM that account for heritable variation in only aboveground biomass due to genotype, provenance, and age cohort. (**E**) Distribution of final predicted elevation of CMEM simulations for scenarios in which aboveground and belowground traits were varied ["ag + bg" from (A)] and for which only aboveground biomass was varied ["ag only" from (D)].

mean and high tides and thus is consequential to how much flooding plants will experience. These differences also result in average vertical accretion rates that could vary more than 1.5-fold [mean vertical accretion rate = 1.45 mm/year (1.18, 1.81)] (Fig. 4B, width of histogram). Predicted rates of carbon accumulation at our sites varied up to 0.32 metric tons C ha^{-1} year^{-1} because of heritable variation [mean C accumulation rate = 0.35 (0.21, 0.53); Fig. 4C, width of histogram], which would lead to estimates of soil carbon storage through 2100 varying by more than twofold in highly organic peat-forming marshes, such as those in the Chesapeake Bay.

Vertical accretion rates were 8% higher and carbon accumulation rates were 18% higher for ancestral cohorts than descendant cohorts. (Fig. 4, B and C). This suggests that modern *S. americanus* marshes may be less resilient to sea level rise and store less carbon compared with marshes from the mid-20th century because of organismal evolution. Across space, changes in plant traits drove soil accretion at Corn Island to be 3% higher than at Sellman Creek and drove soil carbon accumulation rates to be 6% higher (Fig. 4, A to C, green versus gold). Additional evidence from a separate experiment suggests that our estimates of the impact of heritable variation and rapid temporal evolution on accretion and carbon accumulation are robust and possibly conservative (*40*) (figs. S12 and S13 and table S4).

The effect of evolution on ecosystem processes captured in this work is comparable with the effects of rapid environmental change. For example, the magnitude of evolution's influence on vertical accretion that we found is similar to the modeled effect of shifts in salt marsh mineral accretion rates from changes in *S. americanus* stem morphology over 11 years of exposure to elevated CO_2 in a different but related study (*36*). Additionally, by running simulations of the ecosystem model in which we varied the total amount of sea level rise, we

found that the percent difference in carbon accumulation rate between ancestral and descendant genotypes was approximately equal to an additional 4 cm of sea level rise by 2100. The effects of organismal evolution on carbon accumulation are particularly notable given that they would partially offset predicted large increases in future carbon storage in response to global change factors such as increasing atmospheric CO_2 (*41*, *42*).

Much of the projected variation in our predictions of marsh accretion and carbon accumulation was attributable to belowground trait variation (Fig. 4E), which suggests that further study of the effects of genotypic variation on belowground traits can improve forecasts of coastal marshes in which surface accretion is predominantly driven by organic matter accretion. This was revealed by running simulations in which we only accounted for between-genotype variation in aboveground biomass, keeping belowground traits (root-to-shoot ratio and depth of the 95% cumulative root distribution) constant, which aligns with the simplifying assumption that belowground traits covary with aboveground traits (Fig. 1D). Failing to account for heritable variation in belowground traits beyond the variation in aboveground biomass dramatically decreased the predicted uncertainty for accretion and carbon sequestration (Fig. 4, A, D, and E). For example, the variance in final predicted marsh elevation decreased by 68% when only variance in aboveground biomass was included in the simulations (Fig. 4E). Accounting for heritable belowground trait variation also altered the average predictions for each age cohort and provenance. Notably, differences in ecosystem outcomes between age cohorts were larger than those between provenances when belowground trait variation was considered (Fig. 4A), but the opposite was true when only aboveground trait variation was considered (Fig. 4D). However, although belowground trait variation drove variation in ecosystem processes in a highly organic marsh in the Chesapeake Bay, heritable variation and evolution of aboveground plant traits may play a larger relative role in marshes in which accretion rates are driven by mineral sediment capture aboveground, a process mediated by stem density and morphology (*36*).

Observed shifts in plant traits in long-term studies of coastal marshes have been previously thought to reflect plastic responses induced by exposure to environmental pressures (Fig. 1A). For example, there is evidence that *S. americanus* morphology has changed in response to elevated CO_2 and increased nitrogen deposition over the course of 30 years, with putatively plastic trait changes having substantial consequences for model predictions of aboveground sediment capture rate (*36*). Our findings, along with additional evidence of the adaptive capacity of *S. americanus* (*14*, *30*), offer an updated perspective suggesting that plants can evolve at a pace and magnitude that feeds back on ecosystem-level processes (*6*). Failure to account for heritable variation and rapid evolutionary change in ecosystem models (*43*, *44*) might therefore mischaracterize the role that organismal response plays in ecosystem resilience to environmental change that could systematically alter ecosystem-level predictions. For example, our results suggest that failing to account for decadal-scale evolutionary change may overestimate the potential for coastal marshes to build elevation and store carbon. It is increasingly apparent that organismal evolution and ecosystem development occur on similar time scales, which can elicit feedbacks (*6*). Integrative approaches will thus be increasingly important as anthropogenic change continues to challenge our ability to forecast the resilience of at-risk ecosystems such as coastal marshes (*12*).

REFERENCES AND NOTES

1. J. Chacón-Labella *et al*., *Trends Ecol. Evol.* (2022).
2. J. P. Grime, *Plant Strategies, Vegetation Processes, and Ecosystem Properties* (John Wiley & Sons, 2006).
3. S. Díaz, M. Cabido, *Trends Ecol. Evol.* **16**, 646–655 (2001).
4. S. Des Roches *et al*., *Nat. Ecol. Evol.* **2**, 57–64 (2018).
5. N. G. Hairston Jr., S. P. Ellner, M. A. Geber, T. Yoshida, J. A. Fox, *Ecol. Lett.* **8**, 1114–1127 (2005).
6. J. G. Monroe *et al*., *Trends Ecol. Evol.* **33**, 213–225 (2018).
7. T. G. Whitham *et al*., *Nat. Rev. Genet.* **7**, 510–523 (2006).
8. L. A. Donovan, C. M. Mason, A. W. Bowsher, E. W. Goolsby, C. D. A. Ishibashi, *J. Ecol.* **102**, 302–314 (2014).
9. B. Matthews *et al*., *Ecol. Lett.* **14**, 690–701 (2011).
10. J. Pastor, *Ecosystems* **20**, 245–252 (2017).
11. C. Loehle, J. H. K. Pechmann, *Am. Nat.* **132**, 884–899 (1988).
12. T. J. Mozdzer, M. K. McCormick, I. J. Slette, M. J. Blum, J. P. Megonigal, *Sci. Total Environ.* **853**, 157846 (2022).
13. L. J. Weider, P. D. Jeyasingh, D. Frisch, *Evol. Appl.* **11**, 3–10 (2017).
14. M. J. Blum *et al*., *Evol. Lett.* **5**, 422–431 (2021).
15. M. L. Kirwan, J. P. Megonigal, *Nature* **504**, 53–60 (2013).
16. F. Wang *et al*., *Natl. Sci. Rev.* **8**, nwaa296 (2020).
17. J. R. Holmquist *et al*., *Environ. Res. Lett.* **13**, 115005 (2018).
18. J. T. Morris, P. V. Sundareshwar, C. T. Nietch, B. Kjerfve, D. R. Cahoon, *Ecology* **83**, 2869–2877 (2002).
19. S. Fagherazzi *et al*., *Rev. Geophys.* **50**, RG1002 (2012).
20. T. B. H. Reusch, A. R. Hughes, *Estuaries Coasts* **29**, 159–164 (2006).
21. J. T. Morris *et al*., *Earths Future* **4**, 110–121 (2016).
22. G. L. Noyce, A. J. Smith, M. L. Kirwan, R. L. Rich, J. P. Megonigal, *Nat. Geosci.* 10.1038/s41561-022-01070-6 (2023).
23. G. Ottaviani *et al*., *Trends Ecol. Evol.* **35**, 763–766 (2020).
24. G. T. Freschet *et al*., *New Phytol.* **232**, 1123–1158 (2021).
25. F. S. Chapin III, K. Autumn, F. Pugnaire, *Am. Nat.* **142**, S78–S92 (1993).
26. S. Aikio, M. Mari Markkola, *Evol. Ecol.* **16**, 67–76 (2002).
27. Q. Chen, Y. Wang, C. B. Zou, Z.-L. Wang, *Wetlands* **37**, 985–995 (2017).
28. C. P. Carmona *et al*., *Nature* **597**, 683–687 (2021).
29. J. R. Holmquist *et al*., Zenodo (2022).
30. J. L. Summers, B. Bernik, C. J. Saunders, J. S. McLachlan, M. J. Blum, *Evol. Appl.* **11**, 1715–1731 (2018).
31. M. L. Vahsen *et al*., *Evol. Appl.* **14**, 2831–2847 (2021).
32. C. Y. Lumibao, L. Torres Martínez, J. P. Megonigal, S. A. Van Bael, M. J. Blum, *Mol. Ecol.* **31**, 4571–4585 (2022).
33. D. C. Fisher, M. Oppenheimer, *Ambio* **20**, 102–108 (1991).
34. T. J. Mozdzer, J. A. Langley, P. Mueller, J. P. Megonigal, *Biol. Invasions* **18**, 2619–2631 (2016).
35. L. A. Deegan *et al*., *Nature* **490**, 388–392 (2012).
36. M. Lu, E. R. Herbert, J. A. Langley, M. L. Kirwan, J. P. Megonigal, *Nat. Clim. Chang.* **9**, 764–768 (2019).
37. A. R. Hughes, B. D. Inouye, M. T. J. Johnson, N. Underwood, M. Vellend, *Ecol. Lett.* **11**, 609–623 (2008).
38. K. M. Crawford, K. D. Whitney, *Mol. Ecol.* **19**, 1253–1263 (2010).
39. J. A. Schweitzer *et al*., in *Trait-Mediated Indirect Interactions: Ecological and Evolutionary Perspectives*, T. Ohgushi, O.J. Schmitz, R.D. Holt, Eds. (Cambridge University Press, 2012), pp. 371–390.
40. See the Supplementary Materials.
41. A. J. Rietl, J. P. Megonigal, E. R. Herbert, M. L. Kirwan, *Geophys. Res. Lett.* **48**, e2020GL092051 (2021).
42. M. A. Pastore, J. P. Megonigal, J. A. Langley, *Biogeochemistry* **133**, 73–87 (2017).
43. M. Schuerch *et al*., *Nature* **561**, 231–234 (2018).
44. M. L. Kirwan, S. Temmerman, E. E. Skeehan, G. R. Guntenspergen, S. Fagherazzi, *Nat. Clim. Chang.* **6**, 253–260 (2016).
45. M. R. Gale, D. F. Grigal, *Can. J. For. Res.* **17**, 829–834 (1987).
46. M. L. Vahsen *et al*., Zenodo (2022).

ACKNOWLEDGMENTS

We are grateful for the support provided by technicians and volunteers at the Smithsonian Environmental Research Center. In particular, we thank G. Peresta and A. Peresta for assisting with the experimental setup and W. Vahsen for assisting with data collection. We also thank undergraduate students at the University of Notre Dame (E. Ackerley, A. Appling, B. Brown, E. Nguyen, L. Onken, and C. Powers) for assisting with seed germination and root washing, as well as graduate students H. Kleiner and J. Summers for helping with root washing. Last, we thank J. T. Morris for his efforts in transitioning the Cohort Marsh Equilibrium Model (CMEM) into an R package. **Funding:** This work was supported by National Science Foundation DEB 1655781, DEB 1655702 (M.J.B., J.P.M., J.S.M.); National Science Foundation DEB 09050080, DEB 1457100, DEB 1557009 (J.P.M.); United States Coastal Research Program W912HZ-2020003 (M.L.V.); and the Coastal Carbon Research Coordination Network DEB-1655622 (J.R.H.). **Author contributions:** Conceptualization: M.L.V., M.J.B., J.P.M., and J.S.M. Methodology: M.L.V., B.S., and J.S.M. Software development: J.R.H. and K.E.O.T-B. Analysis: M.L.V. and S.J.E. Funding acquisition: M.L.V., M.J.B., J.P.M., and J.S.M.; Writing – original draft: M.L.V. and J.S.M. Writing – review and editing: M.L.V., M.J.B., J.P.M., S.J.E., J.R.H., B.S., K.E.O.T-B., and J.S.M. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** All data and code are available via Zenodo (*46*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq0595
Materials and Methods
Figs. S1 to S13
Tables S1 to S4
References (*47*–*70*)

Submitted 22 March 2022; resubmitted 4 October 2022
Accepted 22 December 2022
10.1126/science.abq0595

SOLAR CELLS

Highly efficient p-i-n perovskite solar cells that endure temperature variations

Guixiang Li[1]†, Zhenhuang Su[2]†, Laura Canil[1], Declan Hughes[3], Mahmoud H. Aldamasy[1], Janardan Dagar[1], Sergei Trofimov[1], Luyao Wang[1]*, Weiwei Zuo[4], José J. Jerónimo-Rendon[4], Mahdi Malekshahi Byranvand[4,5], Chenyue Wang[2], Rui Zhu[6], Zuhong Zhang[6], Feng Yang[6], Giuseppe Nasti[7], Boris Naydenov[1], Wing C. Tsoi[3], Zhe Li[8], Xingyu Gao[2], Zhaokui Wang[9], Yu Jia[6], Eva Unger[1], Michael Saliba[4,5], Meng Li[1,6,8]*, Antonio Abate[1,7]*

Daily temperature variations induce phase transitions and lattice strains in halide perovskites, challenging their stability in solar cells. We stabilized the perovskite black phase and improved solar cell performance using the ordered dipolar structure of β-poly(1,1-difluoroethylene) to control perovskite film crystallization and energy alignment. We demonstrated p-i-n perovskite solar cells with a record power conversion efficiency of 24.6% over 18 square millimeters and 23.1% over 1 square centimeter, which retained 96 and 88% of the efficiency after 1000 hours of 1-sun maximum power point tracking at 25° and 75°C, respectively. Devices under rapid thermal cycling between −60° and +80°C showed no sign of fatigue, demonstrating the impact of the ordered dipolar structure on the operational stability of perovskite solar cells.

The highest power conversion efficiencies (PCEs) of >25% reported for single-junction perovskite solar cells (PSCs) rely on regular n-i-p architectures (*1*). However, inverted p-i-n PSCs have several advantages, including low-temperature processability and long-term operational stability derived from non-doped hole-transporting materials (*2*, *3*). Nonetheless, they have lower PCEs, with only a few certified values exceeding 23% and a bottleneck value of 24% over 10 square millimeter cells (*4*–*6*). This lower performance is mainly correlated with nonradiative recombination losses and reduced charge extraction that stem from the high density of defects in the perovskite bulk and interfacial contacts (*7*, *8*).

[1]Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, Hahn-Meitner-Platz 1, 14109 Berlin, Germany. [2]Shanghai Synchrotron Radiation Facility (SSRF), Shanghai Advanced Research Institute, Chinese Academy of Sciences, 239 Zhangheng Road, Shanghai 201204, China. [3]SPECIFIC, Department of Materials Science and Engineering, Faculty of Science and Engineering, Swansea University, Bay Campus, Fabian Way, Swansea SA1 8EN, UK. [4]Institute for Photovoltaics (*ipv*), University of Stuttgart, Pfaffenwaldring 47, D-70569 Stuttgart, Germany. [5]Helmholtz Young Investigator Group FRONTRUNNER, IEK5-Photovoltaics, Forschungszentrum Jülich, 52425 Jülich, Germany. [6]Key Lab for Special Functional Materials of Ministry of Education, National & Local Joint Engineering Research Center for High-efficiency Display and Lighting Technology, School of Materials Science and Engineering, Collaborative Innovation Center of Nano Functional Materials and Applications, Henan University, Kaifeng 475004, China. [7]Department of Chemical, Materials and Production Engineering, University of Naples Federico II. Naples, pzz.le Vincenzo Tecchio 80, 80125 Naples, Italy. [8]School of Engineering and Materials Science (SEMS), Queen Mary University of London, London E1 4NS, UK. [9]Institute of Functional Nano & Soft Materials (FUNSOM), Jiangsu Key Laboratory for Carbon-Based Functional Materials & Devices, Soochow University, Suzhou 215123, China.
***Corresponding author. Email: mengli@henu.edu.cn (M.L.); luyao.wang@helmholtz-berlin.de (L.W.); antonio.abate@helmholtz-berlin.de (A.A.)**
†These authors contributed equally.

For practical applications, ambient temperature variations can limit PSC performance (*9*) because the perovskite can undergo severe ion migration, phase transition, and temperature-induced strain, leading to lower PCE (*10*–*12*). Cycling over variable temperatures demands that the perovskite tolerate alternating tension and compression in the device structure (*13*). Thus, developing high-efficiency PSCs with thermal cycle stability is critical to advancing PSC application.

We use polymer dipoles to optimize triple cation halide perovskite $Cs_{0.05}(FA_{0.98}MA_{0.02})_{0.95}Pb(I_{0.98}Br_{0.02})_3$ films from bulk to the surface. The polymer dipole promoted the growth of a low-defect crystalline film by reducing the formation energy of the black photoactive phase. The formation of dipoles at the perovskite surface suppressed ion migration and facilitated interfacial charge extraction while enhancing hydrophobicity. We achieved a certified PCE of 24.2% over an active area of 9.6 mm^2, and lab recorded PCEs of 24.6% over 18 mm^2 and 23.1% over 1 cm^2. The high PCE was stable under severe thermal cycling (120 cycles, from −60° to +80°C), demonstrating the resiliency of the crystal structure to the temperature-induced strains.

Film formation and characterization

The alternate symmetric hydro and fluorocarbon units along the polymeric backbone of β-poly(1,1-difluoroethylene) (β-pV2F) result in an ordered molecular dipole distribution. After screening a few molecular weights, 0.5 mg/mL β-pV2F of 180,000 molecular weight was used (figs. S1 and S2). The influence of β-pV2F on the film morphology and work function is shown in Fig. 1. From the top view and cross sectional scanning electron microscope (SEM) images (Fig. 1, A to C), we can observe evident voids at the grain boundaries of the control perovskite film, with an average grain size of ~400 nm (fig. S3A). These defects can create shunting paths and nonradiative recombination centers (*14*). β-pV2F enabled a more compact perovskite film with an enlarged grain size of ~480 nm (Fig. 1, D to F, and fig. S3B). A smaller full width at half maximum in the (001) peak of the x-ray diffraction support an enhanced crystallinity in the β-pV2F–treated perovskite film (fig. S4) (*15*). Furthermore, atomic force microscopy images showed that β-pV2F reduced the surface roughness from 54.4 to 41.1 nm (fig. S5), which is expected to ameliorate coverage with charge-transporting layers (*16*).

Because of the electron-withdrawing effect of the fluorine atoms, the neighboring hydrogen atoms have a partial positive charge density. Then, the all-trans planar zigzag conformation of β-pV2F makes it resemble a Lewis acid, which can interact with the surface of the perovskite (*17*, *18*). Fourier transform infrared spectroscopy revealed that the $-CH_2$ stretching vibration peak shifted from 3025 cm^{-1} of β-pV2F to 3019 cm^{-1} in contact with the target perovskite (fig. S6), suggesting a solid C-H···X dipole interaction between $-CH_2$ moieties and halide ions of the $[PbX_6]^{4-}$ frame. Such polar interaction with the precursors of the perovskite influences the crystallization during film formation and leads to an upward shift of the surface work function (WF) after film formation (Fig. 1G) (*17*, *18*); Fig. 1H displays the increase in WF up to 300 meV for the target perovskite film, which facilitates the interfacial charge extraction and enhances the device's stability (*19*).

The WF shift was near that of standard perovskite film treated with β-pV2F only at the surface, which indicated that as the crystal growth proceeded, β-pV2F was partially expelled from the bulk and assembled on the perovskite surface (*2*, *20*). The fluorine-exposed surface arrangement induced hydrophobicity (figs. S7 and S8). We measured reduced nonradiative recombination and improved interfacial charge transfer in target perovskites (figs. S9 to S13) (*21*–*24*). This scenario is expected to enhance the solar cells' efficiency and stability (*25*, *26*).

To acquire an in-depth perspective on the promoted perovskite crystallization kinetics, we performed synchrotron-based in situ grazing incidence wide-angle x-ray scattering (GIWAXS) measurements to monitor the entire film formation process (several different stages are shown in Fig. 2, A and B). The initial t_1 stage (during the first 25 s) revealed the scattering halo at scattering vector q values from 8 to 8.5 nm^{-1} from the solvated colloidal sol precursor. The signal transition at t_2 (25 s) originated from dripping antisolvent, where the rapid solvent extraction caused the disappearance of the diffraction signal. Subsequently,

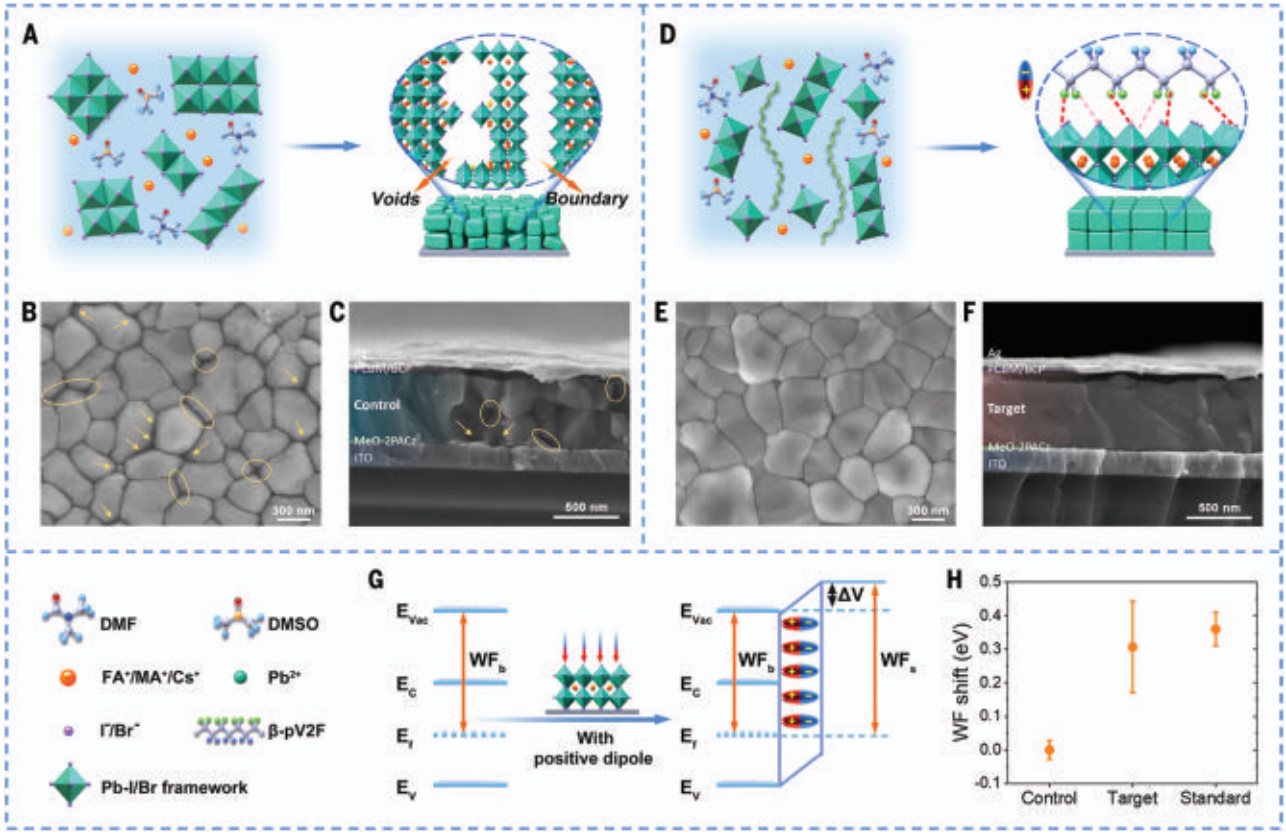


Fig. 1. Working mechanism and morphology characterization of perovskite films. Schematic of processing (**A**) control and (**D**) target perovskites. (**B**) Top view and (**C**) cross sectional SEM images of control perovskites. (**E**) Top view and (**F**) cross sectional SEM images of target perovskites. (**G** and **H**) WF shift related to perovskite functionalized with β-pV2F.

the spin coating process was performed at stage t_3, where supersaturated solvate intermediate emerges. The t_4 near 60 s represented annealing staging. Stage t_5 revealed the intermediate phase signal with annealing. Stage t_6 was the perovskite evolution process. Stage t_7 described the cessation of further crystal growth.

Comparing GIWAXS patterns (Fig. 2, A and B), the weakened diffraction signal in the initial 60 s suggested that the initial solvated phase of DMSO-DMF-PbX_2—in which DMSO is dimethylsulfoxide, DMF is dimethylformamide, and X is a halide (I^-, Br^-)—was suppressed. This effect could be ascribed to the initial solvated phase isolated by the long-chain β-pV2F molecules (*27*). The intermediate phase concentration was lower in the target than in the control (fig. S14). The scattering feature centered at q = ~10 nm^{-1} along the (001) crystal plane observed in the cast film indicated that the colloid had solidified and converted into a black phase. Notably, we found that the perovskite phase of the target emerged earlier than that of the control ($\Delta t_t > \Delta t_c$), which implied that β-pV2F promoted the conversion of the intermediate phase to the perovskite black phase. The fast phase inversions were

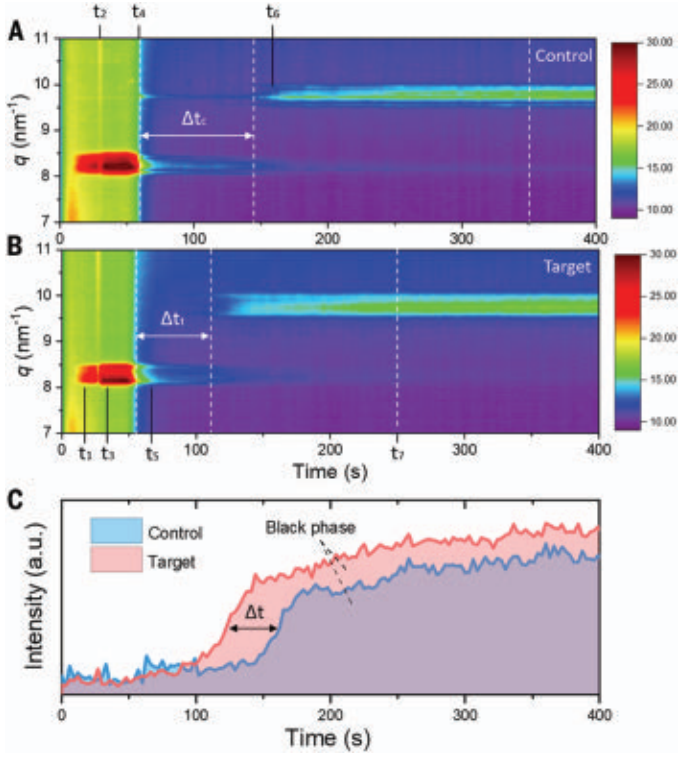


Fig. 2. Crystallization kinetics of perovskite films. In situ GIWAXS spectra during forming (**A**) control and (**B**) target perovskite films. (**C**) Time-resolved integrated peak area intensity for black phases of control and target perovskites.

associated with the lower formation energy (*28*, *29*) and could be attributed to β-pV2F rapidly aggregating dispersed PbX_2 and organic salts during the elimination of DMSO and DMF (*30*).

Moreover, the target ceased crystal growth sooner at 250 s than the control at 350 s. When the crystallization is completed (stage t_7), the signal is more intense in the target than in the control (Fig. 2C). This result indicates that the target perovskite film is more ordered. The time-dependent in situ GIWAXS intensity profiles with other scattering vectors, such as q = ~20 nm^{-1} corresponding to the (002) crystal plane (fig. S15), showed the same phase transition trend. Thus, β-pV2F controls the perovskite crystallization kinetics by lowering the perovskite formation energy, promoting phase conversion, and enabling a more ordered crystal structure (fig. S16).

Photovoltaic performance

The photovoltaic performance of inverted p-i-n PSCs with control and polymer-modified perovskite films is shown in Fig. 3. The device architecture is glass/indium tin oxide (ITO)/self-assembled [2-(3,6-dimethoxy-9H-carbazol-9-yl)ethyl]phosphonic acid (MeO-2PACz)/perovskite/ [6,6]-phenyl-C61-butyric acid methyl ester ($PC_{61}BM$)/bathocuproine (BCP)/silver (Ag) (fig. S17). Typical current-voltage (*J-V*) curves for the PSCs (Fig. 3A) were measured with a device area of 18 mm^2. The control PSCs had a PCE of 22.3%, with short-circuit current density (J_{sc}) of 24.7 mA/cm^2, V_{oc} of 1.13 V, and fill factor (FF) of 80.2%. With β-pV2F, the device performance improved with a V_{oc} of 1.18 V, a J_{sc} of 24.8 mA/cm^2, and a FF of 84.3% for a PCE of 24.6%. The target PSCs' reverse- and forward-sweep *J-V* curves (fig. S18) had negligible hysteresis. A PCE of 24.2% for an aperture area of 9.6 mm^2 was obtained from the independent accredited certification institute of Test and Calibration Center of New Energy Device and Module, Shanghai Institute of Microsystem and Information Technology (SIMIT), Chinese Academy of Sciences (fig. S19). We also recorded a PCE of 23.1% for devices with a working area of 1 cm^2 (Fig. 3B).

From the external quantum efficiency (EQE) spectra (Fig. 3C), we calculated an integrated J_{sc} of 24.3 and 24.4 mA/cm^2 for control and target devices, respectively, comparable to the values extracted from the *J-V* curves. The optical bandgaps of both perovskite absorbers determined by the x axis intercept of the EQE linear are shown in fig. S20 (*31*, *32*). The statistical distribution of the device parameters collected from 38 devices shows an improved PV performance and increased reproducibility with β-pV2F (fig. S21) (*33*), which we explain with better charge extraction and reduced nonradiative recombination (*27*).

The stabilized power outputs at the maximum power point (MPP) are plotted in Fig. 3D. The control device showed a persistent attenuation in efficiency under continuous 1 equivalent sun illumination for 400 s. By contrast, the tracked target device yielded highly stable power output and even progressively improved performance, which we attributed to the light-soaking effect (*34*). The stability of unencapsulated devices under working conditions shows that target PSCs retain 96% of the initial PCE after continuous MPP tracking for 1000 hours. By contrast, control PSCs decay to 84% of their original PCE (Fig. 3E). Device stability statistics (n = 12) are presented in fig. S22. After heating the device to 75°C, 88% PCE was retained in the target device in contrast to only 56% in the control (fig. S23).

We further evaluated device stability against temperature variations. The *J-V* curves in fig. S24 show that the control device's PV parameters (J_{sc}, FF, and V_{oc}) exhibited large fluctuations when tested at temperatures ranging from −60° to +80°C (table S1). However, this variation was suppressed in the target device (fig. S25 and table S2). Furthermore, compared with the control, the target device had reduced hysteresis and its hysteresis factor was relatively stable under temperature variations (fig. S26). The statistical PCE distribution in Fig. 4, A and B, indicates that the β-pV2F stabilization effect is highly reproducible. All performance parameter evolution is reported in figs. S27 and S28. Subsequently, the unencapsulated devices were aged under rapid thermal cycling (TC) between −60° and +80°C, swept at a rate of 20°C per minute. As shown in Fig. 4, C and D, the control device suffered a severe decline of 75.6% at +80°C and 63.0% at −60°C, whereas the target device retained 93.9% at 80°C and 88.7% at −60°C of its initial value after 120 thermal cycles.

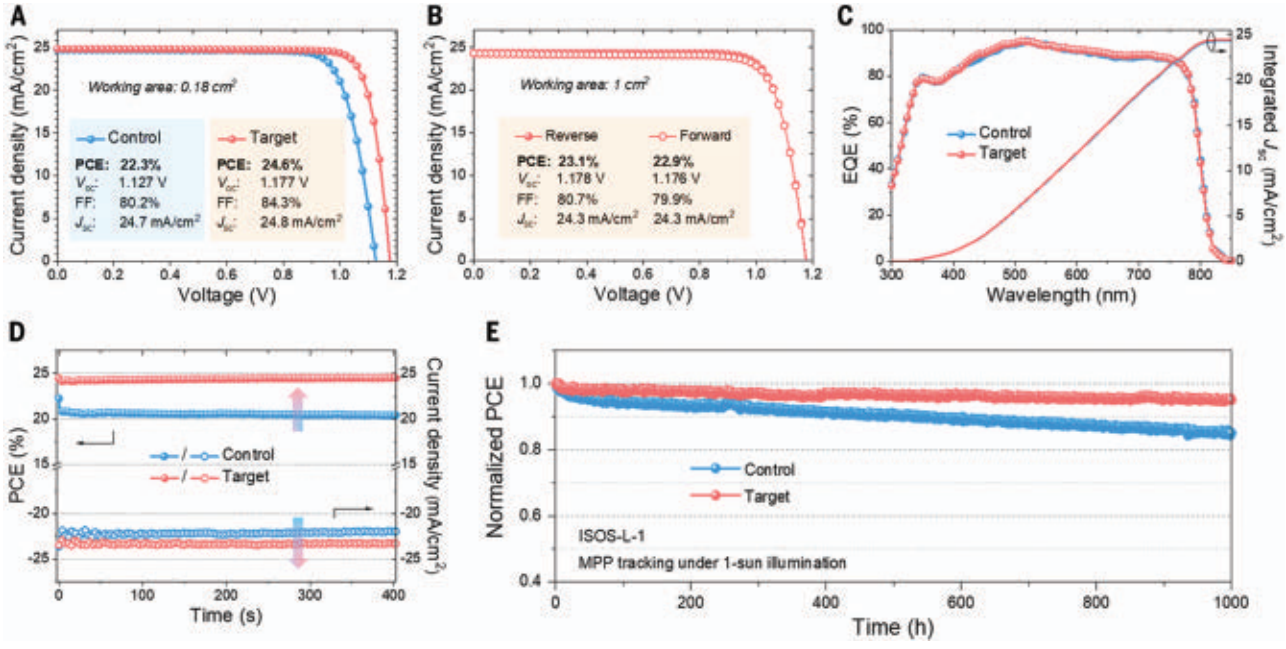


Fig. 3. PV performance of perovskite solar cells. (**A**) *J-V* curves of control and target PSCs under a device area of 0.18 cm^2. (**B**) *J-V* curves with the reverse and forward sweeps for large-area target PSCs (1 cm^2). (**C**) EQE spectra and integrated J_{sc} for control and target PSCs. (**D**) Stabilized power outputs with evolving current density at the maximum power points as a function of time for the best-performing PSCs. (**E**) Long-term stability at maximum power point tracking under room-temperature continuous illumination in N_2 atmosphere for unencapsulated PSCs (ISOS-L-1 procedure).

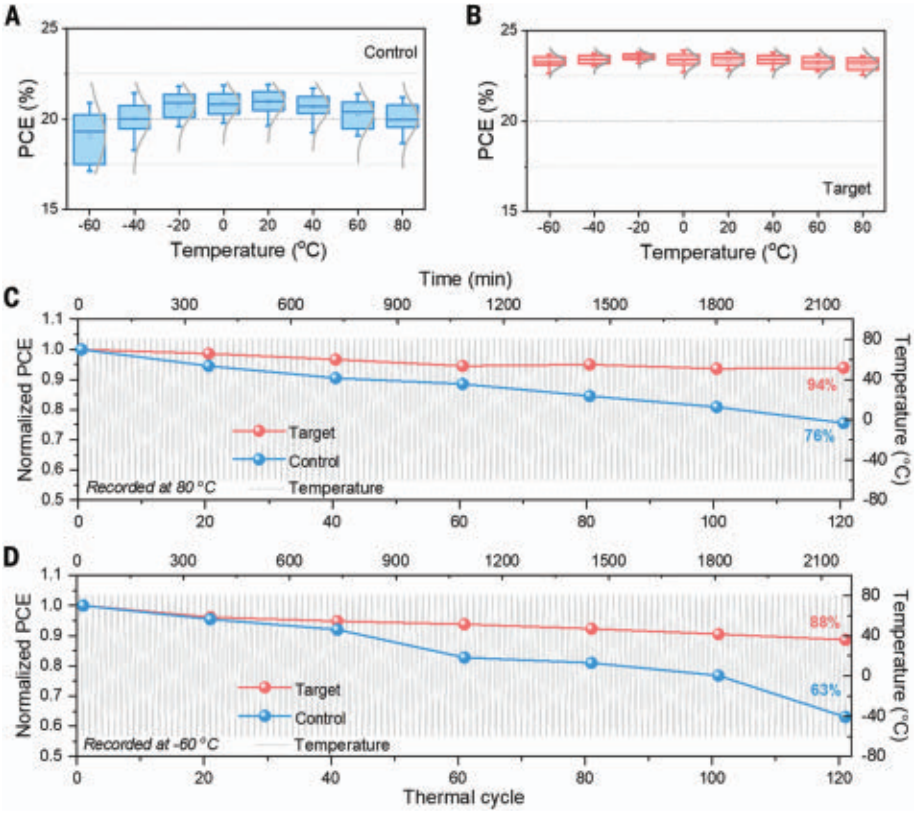


Fig. 4. Thermal cycling stability of perovskite solar cells. Statistical temperature dependence PCE profiles of (**A**) control PSCs and (**B**) target PSCs. PCE evolution recorded at (**C**) +80°C and (**D**) −60°C of control and target PSCs against thermal cycles between −60° and +80°C. The rapid thermal cycling was implemented with a ramp rate of 20°C per minute. There is an extra 2-minute waiting window for the device to reach thermal equilibrium when cycling to −60° and +80°C. The temperature starts at room temperature and is heated to +80°C, then cooled to −60°C. The progress ends at room temperature. The time per complete cycle is 18 minutes.

Device film morphologies and structures during thermal cycling

The difference in device performance stems from use of β-pV2F in the perovskite film. We characterized the morphology and crystal structure of perovskite films undergoing aging with thermal cycling to identify the impact of β-pV2F. The film aging followed the same device protocol: 120 rapid thermal cycles between −60° and +80°C at a rate of 20°C per minute. As observed from the SEM images in fig. S29, control films exhibited severe morphological degradation with enlarged grain boundaries and voids. Such degradation features were not detected in the aged target films (fig. S30), which appeared nearly identical to the pristine film shown in Fig. 1E. This result indicates that the temperature-induced degradation of perovskite films is suppressed in target perovskites (fig. S31) (*35*).

We observed additional GIWAXS peaks forming in the control perovskite after three thermal cycles (Fig. 5A). Specifically, in the second cycle, the peak for PbI_2, a degradation product, emerged at q = 9.2 nm^{-1}. During the third thermal cycle, additional peaks ~8.2 and 8.6 nm^{-1} formed, corresponding to the hexagonal photoinactive polytypes 4H and 6H from $Cs_{0.05}(FA_{0.98}MA_{0.02})_{0.95}Pb(I_{0.98}Br_{0.02})_3$ perovskite (fig. S32A) (*36*, *37*). This result indicates that the control perovskite undergoes irreversible phase changes. The generation of these phases may originate from the lattice deformation at the grain boundaries caused by the mutual extrusion of unit cells from neighbored crystals of different orientations (*36*). Such phenomena were not observed in the target perovskite (Fig. 5B), indicating high structural stability (fig. S32B).

For q values from 16 to 19 nm^{-1}, we observed additional peaks in both control and target (figs. S33 and S34), corresponding to the tetragonal phase (β phase) (*10*, *38*). The tetragonal phase was only retained in the cold temperature region. This suggests that the degradation products of perovskite under thermal cycling include irreversible PbI_2, 4H, and 6H, and reversible tetragonal phase transition, jointly contributing to device performance degradation. Temperature-resolved azimuthally integrated intensity patterns (figs. S35 and S36) indicate that β-pV2F suppresses the phase transitions. We found that suppressing the phase transition also suppressed ion migration in the complete device under working conditions, i.e., lower hysteresis (fig. S37).

Because of differences in thermal expansion coefficients between the perovskite film and the substrate, temperature variation induces strain in the perovskite (*39*, *40*). The control perovskite underwent substantial lattice strain evolution (−0.13 to 0.57%) during thermal cycling (Fig. 5C). We observed that the perovskite strain drifted with temperature cycling, showing a constant lattice parameter change in perovskite. By contrast, the target perovskite exhibits stable strain cycling in a narrower range (−0.06 to 0.38%), corresponding to a recoverable crystal structure and releasable lattice strain (tables S3 to S5). We propose that a strain-buffering and lattice-stabilizing effect exists in target perovskite because β-pV2F creates a self-assembly polymeric layer that coats the crystals within the perovskite film reducing friction during thermal cycling (figs. S38 and S39) (*41*–*44*).

Conclusions

Thermal stress experienced in normal working conditions induces phase transitions and lattice strains that hamper the stability of perovskite solar cells (PSCs). Coating the crystals comprising the perovskite film with polymer dipoles results in a strain-buffering and lattice-stabilizing effect that mitigates the impact of thermal stress. We selected the specific polymer dipole β-poly(1,1-difluoroethylene) (β-pV2F). The β-pV2F highly ordered dipolar structure interacts with specific perovskite components enabling control of perovskite film crystallization during the processing and energy alignment with the charge-selective contacts within the device. We reported β-pV2F devices with improved power conversion efficiency up to 24.6% on an active area of 18 mm^2 and 23.1% over a larger area of 1 cm^2 (certified PCE of 24.24% with an active area of 9.6 mm^2 from SIMIT). The β-pV2F strain-buffering effects enabled stable power output at temperatures as high as 75°C and rapid temperature variation

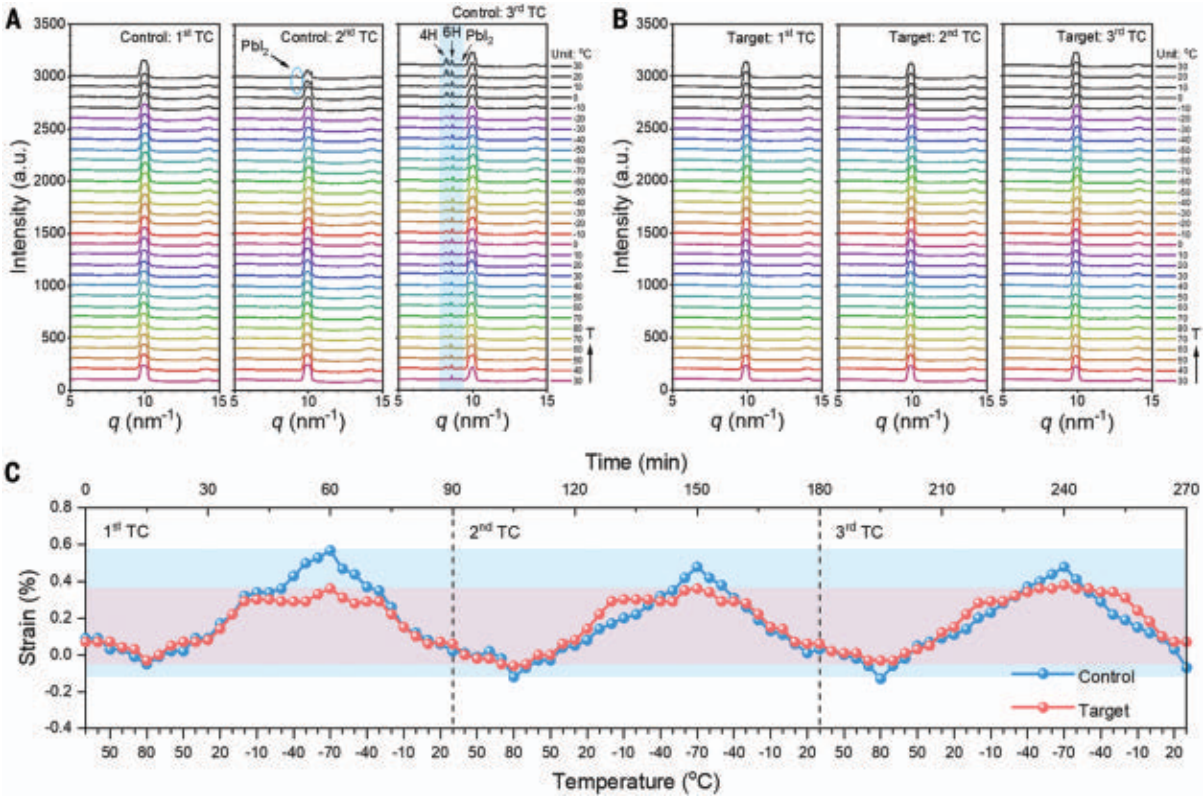


Fig. 5. Perovskite structural evolution during temperature cycling. The temperature-resolved GIWAXS profiles for (**A**) control and (**B**) target perovskites. (**C**) The temperature-resolved lattice strain for control and target perovskites. (The temperature starts from room temperature, heating to +80°C and then cooling to −60°C. The progress ends at room temperature. The time per complete cycle is 90 minutes.)

between −60° and +80°C. Our work identifies a new strategy for making stable perovskite solar cells.

REFERENCES AND NOTES

1. The National Renewable Energy Laboratory, US Department of Energy, Office of Energy Efficiency and Renewable Energy, Alliance for Sustainable Energy LLC., Best Research-Cell Efficiency Chart, https://www.nrel.gov/pv/cell-efficiency.html.
2. X. Zheng *et al*., *Nat. Energy* **5**, 131–140 (2020).
3. Y.-N. Lu *et al*., *Energy Environ. Sci.* **14**, 4048–4058 (2021).
4. Z. Li *et al*., *Science* **376**, 416–420 (2022).
5. H. Chen *et al*., *Nat. Photonics* **16**, 352–358 (2022).
6. X. Li *et al*., *Science* **375**, 434–437 (2022).
7. M. Degani *et al*., *Sci. Adv.* **7**, eabj7930 (2021).
8. D. Luo *et al*., *Science* **360**, 1442–1446 (2018).
9. W. Tress *et al*., *Nat. Energy* **4**, 568–574 (2019).
10. S. Kahmann *et al*., *ACS Appl. Mater. Interfaces* **14**, 34253–34261 (2022).
11. X. Zhang, S.-H. Wei, *Phys. Rev. Lett.* **128**, 136401 (2022).
12. D.-J. Xue *et al*., *Nat. Commun.* **11**, 1514 (2020).
13. R. Cheacharoen *et al*., *Energy Environ. Sci.* **11**, 144–150 (2018).
14. C. Sun *et al*., *Nat. Commun.* **12**, 2207 (2021).
15. Z. Wang *et al*., *Nat. Energy* **2**, 17135 (2017).
16. L. Xu *et al*., *Adv. Mater.* **34**, 2107111 (2022).
17. L. Canil *et al*., *Energy Environ. Sci.* **14**, 1429–1438 (2021).
18. A. J. de Jesus Silva, M. M. Contreras, C. R. Nascimento, M. F. da Costa, *Heliyon* **6**, e04573 (2020).
19. P. Metrangolo, L. Canil, A. Abate, G. Terraneo, G. Cavallo, *Angew. Chem. Int. Ed.* **61**, e202114793 (2022).
20. Y. Moriya, R. Ishikawa, S. Akiyama, K. Ueno, H. Shirai, *Chem. Lett.* **49**, 87–90 (2019).
21. J. Jeong *et al*., *Nature* **592**, 381–385 (2021).
22. G. Li *et al*., *Adv. Energy Mater.* **11**, 2101539 (2021).
23. D. Koo *et al*., *Adv. Energy Mater.* **10**, 2001920 (2020).
24. S. Chen *et al*., *Science* **373**, 902–907 (2021).
25. M. Stolterfoht *et al*., *Adv. Mater.* **32**, e2000080 (2020).
26. L. Canil *et al*., *Adv. Energy Mater.* **11**, 2101553 (2021).
27. Y. Li *et al*., *Joule* **6**, 676–689 (2022).
28. Y. Cho *et al*., *Adv. Energy Mater.* **8**, 1703392 (2018).
29. H. Hu *et al*., *J. Mater. Chem. A Mater. Energy Sustain.* **8**, 1578–1603 (2020).
30. M. Qin *et al*., *Adv. Mater.* **32**, e2004630 (2020).
31. A. Y. Alsalloum *et al*., *Energy Environ. Sci.* **14**, 2263–2268 (2021).
32. P. S. C. Schulze *et al*., *Sol. RRL* **4**, 2000152 (2020).
33. R. Azmi *et al*., *Science* **376**, 73–77 (2022).
34. S. Wu *et al*., *Joule* **4**, 1248–1262 (2020).
35. Z. Dai *et al*., *Science* **372**, 618–622 (2021).
36. M. Qin *et al*., *Adv. Mater.* **31**, e1901284 (2019).
37. P. Gratia *et al*., *ACS Energy Lett.* **2**, 2686–2693 (2017).
38. W. Peng *et al*., *Angew. Chem. Int. Ed.* **55**, 10686–10690 (2016).
39. C. Zhu *et al*., *Nat. Commun.* **10**, 815 (2019).
40. J. Zhao *et al*., *Sci. Adv.* **3**, eaao5616 (2017).
41. L. Wang, Q. Gong, S. Zhan, L. Jiang, Y. Zheng, *Adv. Mater.* **28**, 7729–7735 (2016).
42. L. T. Hieu, S. So, I. T. Kim, J. Hur, *Chem. Eng. J.* **411**, 128584 (2021).
43. W. Chen *et al*., *Adv. Funct. Mater.* **29**, 1808855 (2019).
44. L. Zuo *et al*., *Sci. Adv.* **3**, e1700106 (2017).

ACKNOWLEDGMENTS

The authors acknowledge the support of all the technicians at Helmholtz-Zentrum Berlin (HZB). The authors thank beamline BL14B1 at the Shanghai Synchrotron Radiation Facility (SSRF) for providing the beam time. G.L., L.C., and M.H.A. thank the support from HyPerCells graduate school at HZB. R.Z. was supported by the National Natural Science Foundation of China (22103022). G.L. thanks the Chinese Scholarship Council (CSC) for its financial support (201906150131). M.S. and W.Z. thank the German Research Foundation (DFG) for funding (SPP2196, 431314977/GRK 2642). M.S. acknowledges funding by ProperPhotoMile. Project ProperPhotoMile is supported under the umbrella of SOLAR-ERA.NET Cofund 2 by the Spanish Ministry of Science and Education and the AEI under the project PCI2020-112185 and CDTI project number IDI-20210171; the Federal Ministry for Economic Affairs and Energy based on a decision by the German Bundestag project number FKZ 03EE1070B and FKZ 03EE1070A and the Israel Ministry of Energy with project number 220-11-031. The European Commission supports SOLAR-ERA.NET within the EU Framework Programme for Research and Innovation HORIZON 2020 (Cofund ERA-NET Action, 786483). This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 804519). **Author contributions:** G.L., L.W., M.L., and A.A. conceived the idea. G.L. and M.L. designed the experiments. G.L., Z.S., L.W., S.T., M.H.M., Z.Z., F.Y., and M.L. fabricated and characterized perovskite films and devices. Z.S. and C.W. conducted the GIWAXS measurements. G.L., D.H., L.W., W.Z., R.Z., F.Y., Y.J., and M.L. performed the device performance measurements. G.L., Z.S., L.C., L.W., M.L., and A.A. participated in the data analysis and result discussions. G.L. composed the manuscript. D.H., M.H.M., J.D., S.T., W.Z., J.J.J.-R., M.M.B., G.N., B.N., M.L., and A.A. contributed to suggestions for the manuscript. L.W., W.C.T., Z.L., X.G., Z.W., E.U., M.S., M.L., and A.A. provided expertise and supervised the work. All authors reviewed the manuscript. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

License information: Copyright © 2023 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.add7331
Materials and Methods
Supplementary Text
Figs. S1 to S39
Tables S1 to S5
References (*45*–*52*)

Submitted 30 June 2022; accepted 27 December 2022
10.1126/science.add7331

ScienceCareers online @sciencecareers.org

POSTDOCTORAL FELLOW- CANCER RESEARCH

University of Texas Health Science Center at San Antonio
(UT Health SA)

The Mays Cancer Center, an NCI-designated cancer center at UT Health SA, is recruiting **postdoctoral fellows for the Cancer Research Training Program**, which is funded by an institutional Ruth L. Kirschstein National Research Service Award (T32 CA148724) from NIH/NCI. The Program supports highly motivated individuals for intensive research training in all aspects of cancer biology under the direction of a faculty mentor in the Program. The focus of the research may be basic, translational, or clinical. All trainees will also participate in seminars, workshops, and/or lectures as specified by the Program. Our goal is to provide a quality training to facilitate career advancement.

UT Health SA is a research-intensive institution located in a gateway to the picturesque Texas Hill Country. San Antonio is a vibrant, multicultural city with year-round recreational activities, affordable housing, and a high standard of living.

Eligibility and application information can be found at:
https://lsom.uthscsa.edu/dcsa/education/nci-t32/

Applicants must be either U.S. citizens or U.S. Permanent Resident to be eligible for support on this training grant.

All postdoctoral appointments are designated as security sensitive positions.

UT Health SA is an Equal Opportunity/Affirmative Action Employer. All qualified applicants will receive consideration for acceptance into the program without regard to sex, gender identity, sexual orientation, race, color, religion, national origin, disability, protected Veteran status, age, or any other characteristic protected by law or policy.

FACULTY POSITIONS IN VIROLOGY & IMMUNOLOGY

Department of Microbiology
BU Chobanian & Avedisian School of Medicine

The Department of Microbiology (**www.bumc.bu.edu/microbiology**) invites applications for faculty positions at **Assistant or Associate Professor level**. We invite applicants with broad research interests in virology and in immunology, with an emphasis on virus-host interactions and microbial pathogenesis. The successful applicant will join a collaborative, vibrant, and growing community of researchers in an environment that values creativity, collaboration and social and cultural diversity of our faculty, staff and students. The successful applicant will expand and complement existing departmental strengths in virology, host–pathogen interactions, innate immunity and inflammation. Applicants should have a demonstrated and ongoing record of research excellence with a focus that includes basic or translational studies on investigating virology and/or immune responses to human pathogens. The Boston University Medical Campus includes a number of centers of excellence which offer rich collaborative opportunities, including the Center for Regenerative Medicine, the Center for Network Systems Biology, the National Emerging Infectious Diseases Laboratories, a T32 training grant in Immunology and supports state-of-the-art core facilities. The successful candidate is expected to oversee a vigorous, innovative, and extramurally funded research program, to participate in the training of graduate students and post-doctoral fellows and the teaching activities of the department, and support our institutional commitment to ensure BU is inclusive, equitable, diverse and a place where all constituents can thrive.

To be considered, please submit a curriculum vitae, a summary of research accomplishments and future research goals, brief teaching and diversity statements, and contact information to solicit at least 3 letters of recommendation to Ms. Kathleen Marinelli **(kfurness@bu.edu)**. Applications will be considered as they are received, with the positions to be filled during this academic year, beginning after **March 1, 2023**.

Boston University School of Medicine is an equal opportunity/affirmative action employer, committed to building and maintaining robust diversity within all components of its community, and strongly welcomes applications from women, persons with disabilities, protected veterans, and under-represented minorities.

TENURED OR TENURE TRACK FACULTY POSITION IN MICROBIOMES

INDIANA UNIVERSITY BLOOMINGTON

As part of a major, multi-year initiative to expand research activities across the IU Bloomington campus, the Microbiology program in the Department of Biology at Indiana University Bloomington (IUB) invites applications at the **Associate or Full Professor level** in the area of Microbiomes; established Assistant Professors will also be considered. We welcome applicants who study microbial interactions using integrative approaches and methods, in any relevant environment, from soils to humans. Start date is flexible, from **Aug 2023 to Aug 2024**.

Applicants must hold a Ph.D. in relevant fields, with a strong record of research accomplishments and external funding. Successful candidates will be expected to maintain a vigorous externally funded research program, and to participate in teaching at the undergraduate and graduate levels. The Department of Biology is committed to recruiting and retaining diverse faculty, and we strongly encourage applications from members of historically marginalized groups. We also welcome candidates who have demonstrated a commitment to working with people from groups that are underrepresented in STEM, through teaching, mentoring, or administration.

For more detailed application instructions and conditions, please see: **https://indiana.peopleadmin.com/postings/16243**. Applications will start to be evaluated on **February 15, 2023** but the position will remain open until filled. Please address inquiries concerning the search to Jennifer Tarter at 812-856-3984 or **jenjones@indiana.edu**.

TENURED OR TENURE TRACK FACULTY POSITION IN MODELS OF METABOLIC REGULATION

INDIANA UNIVERSITY BLOOMINGTON

As part of a major, multi-year initiative to expand research activities across the Indiana University Bloomington campus, the Genome, Cell, and Developmental Biology program in the Department of Biology (**https://biology.indiana.edu/graduate/genome-cell-development/index.html**) invites applications at the **Associate or Full Professor level** in the area of Metabolism. Established Assistant Professors will also be considered. We are particularly interested in scientists who are using model organisms to study the regulation of metabolism in development and disease. Scientists studying the molecular mechanisms by which the environment influences metabolism and growth are also encouraged to apply. Start dates of **Fall 2023 or Fall 2024** are possible.

Our vision is to continue building a multi-investigator group that will define IU as a preeminent center for studies of metabolism and metabolomics in model organisms. This effort is part of a significant expansion in the life sciences at IU Bloomington. The successful applicant will join a vibrant group of investigators with related research programs within the Department of Biology and across campus in the Departments of Chemistry, Molecular and Cellular Biochemistry, as well as IU School of Medicine and the School of Public Health.

The selected candidate will be provided with a competitive startup package and salary and will have access to outstanding research resources, including state-of-the-art facilities for genomics and bioinformatics, light and electron microscopy (including cryo-EM), flow cytometry, protein analysis, mass spectrometry, and crystallography. Applicants must hold a Ph.D. in a relevant field, with a strong record of research accomplishments and funding. Successful candidates will be expected to maintain a vigorous externally funded research program and participate in teaching at the undergraduate and graduate levels. The Department of Biology is committed to recruiting and retaining diverse faculty, and we strongly encourage applications from members of historically marginalized groups. We also welcome candidates who have demonstrated a commitment to working with people from groups that are underrepresented in STEM, through teaching, mentoring, or administration.

For more detailed application instructions and conditions, please see **https://indiana.peopleadmin.com/postings/16241**. Applications will start to be evaluated on **February 15, 2023**, but the position will remain open until filled. Please address inquiries concerning the search to Jennifer Tarter at 812-856-3984 or **jenjones@indiana.edu.**

Boston Children's Hospital
Where the world comes for answers

HARVARD MEDICAL SCHOOL
TEACHING HOSPITAL

SENIOR FACULTY MEMBER
LABORATORIES OF COGNITIVE NEUROSCIENCE

The Division of Developmental Medicine Laboratories of Cognitive Neuroscience at Boston Children's Hospital (BCH), one of Harvard Medical School's teaching hospitals, is recruiting a new **senior faculty member**. The Division of Developmental Medicine (DDM) is one of the largest clinical and research divisions focused on the diagnosis and treatment of developmental disorders in the United States. The division provides clinical care for children from infancy through adolescence, conducts research, and trains pediatricians, psychologists and cognitive neuroscientists. The Laboratories of Cognitive Neuroscience (LCN) is headed by Charles A. Nelson III, PhD; the LCN includes Dr. Nelson's laboratory as well as labs led by the following PI's: Susan Faja PhD, Carol Wilkinson MD, PhD and Anne Arnett, PhD.

The new faculty member would join the DDM and establish a lab in the LCN. Researchers at the LCN are dedicated to understanding neural and behavioral development of individuals diagnosed with or at higher likelihood for various developmental disorders, as well as typically developing infants and children. Such risk factors could be either genetic (e.g., single gene disorders, such as TSC) or environmental (such as infants growing up in adverse environments). Researchers in the LCN are experts from a wide range of fields, including psychology, neuroscience, and education, and they work in collaboration with clinical investigators in Developmental-Behavioral Pediatrics in the DDM as well as in other disciplines, including, child neurology, and child psychiatry. Beyond the dedicated resources within the Faja, Arnett, Wilkinson, and Nelson labs, the DDM and LCN also provide shared resources including behavioral testing rooms, technical support and staff to assist with state-of-the-art neuroimaging, neurophysiology, behavioral, and clinical evaluations and remote data collection. LCN faculty benefit from several established, effective recruitment pipelines to support study enrollment and administrative support for grant writing and administration.

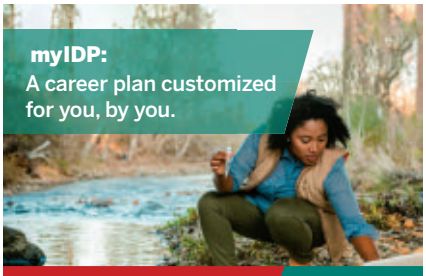
Candidates with methods expertise in MRI, genetics, data science, clinical trials, and speech language pathology are encouraged to apply.

The ideal candidate will have expertise that complements existing research programs, which include:
• Neurodevelopment of autism spectrum disorder
• Biomarkers of ADHD diagnosis and treatment response
• Neuro-genetic etiology of autism spectrum disorder, ADHD, and genetic syndromes like Down Syndrome and Fragile X
• Impact of adversity and early neglect on neurodevelopmental disorders

The candidate must be a PhD, MD, or MD/PhD, and is expected to conduct independent and collaborative research. The candidate will have a track record of independent research funding, preferably from Federal agencies and scholarly productivity as well as successful mentorship of trainees and junior colleagues. Academic appointment at Harvard Medical School will be commensurate with clinical and research experience, training and achievements at the Associate Professor or Professor level.

Please send your CV and a statement of interest to: Search Committee, Attention: Sandra Maislen, Division of Developmental Medicine, BCH 3185, Boston Children's Hospital, 300 Longwood Ave., Boston, MA 02115 or send via email to **sandra.maislen@childrens.harvard.edu**.

We are an equal opportunity employer and all qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability status, protected veteran status, gender identity, sexual orientation, pregnancy and pregnancy-related conditions or any other characteristic protected by law.



There's only one Science

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.

myIDP

Visit the website and start planning today!
myIDP.sciencecareers.org

Science Careers AAAS In partnership with:

UCSF University of California San Francisco
UIC
BURROUGHS WELLCOME FUND
University of Massachusetts Medical School UMASS
FASEB Federation of American Societies for Experimental Biology

utmb Health

FACULTY POSITIONS
UNIVERSITY OF TEXAS MEDICAL BRANCH

The newly established Institute for Drug Discovery (IDD, **https://www.utmb.edu/idd/home**) at the University of Texas Medical Branch (UTMB), Galveston, Texas, seeks outstanding candidates for **five tenure-track positions at the rank of Assistant Professor** in medicinal chemistry, therapeutic antibodies, and nanoformulation. A Ph.D. degree (or equivalent) with postdoctoral (or equivalent) experience is expected.

The applicants will join UTMB's world-class research program with strengths in infectious disease, neuroscience, aging, and other disciplines. In addition to a highly collaborative environment, UTMB offers outstanding research facilities for next-generation sequencing, mass spectrometry, optical microscopy, flow cytometry, solution biophysics, NMR, X-ray crystallography, cryo-electron microscopy, scientific computing, and bioinformatics. Excellent opportunities for scientific interactions exist through sister institutions in the Greater Houston area, Gulf Coast Consortia, and the Keck Center for Interdisciplinary Bioscience.

Through academic and industrial partnerships, UTMB has led the way in discovery of small molecules and biologics for treatment of diseases. The IDD seeks to build on this success for further therapeutics development and has invested in brand-new synthetic chemistry laboratories with state-of-the-art fume hoods to achieve this goal. Our established drug discovery for pandemic preparedness program is powered by the Galveston National Laboratory, the premier biosafety-level-4 facility in the world, and well-established pathogen models at UTMB. Minorities and members of underrepresented groups are encouraged to apply. The first application review will consider applications submitted by **March 20, 2023**, with additional reviews quarterly until positions are filled. Applications will be accepted with additional reviews quarterly until **December 2023**.

An attractive recruitment package of competitive salary, generous start-up funding, new or renovated laboratory space will be offered. Interested applicants should submit a single PDF (<10 MB) containing the following: **1)** cover letter, **2)** curriculum vitae, **3)** summary of research accomplishments and future goals (2-3 pages), **4)** diversity statement which follows the NIH PEDP guidelines, **5)** teaching and mentoring statement (1 page) and **6)** contact information for three references. The combined PDF and any inquiries should be sent by email to: Valerie Rodecap at **utmbidd@utmb.edu**.

UTMB Health strives to provide equal opportunity employment without regard to race, color, national origin, sex, age, religion, disability, sexual orientation, gender identity or expression, genetic information or veteran status. As a Federal Contractor, UTMB Health takes affirmative action to hire and advance women, minorities, protected veterans and individuals with disabilities.

By **Hasan Merali**

Finding time for my passion

For years, I have wanted to write a book about the lessons adults can learn from children, drawing on my work as a pediatric emergency medicine physician and child health researcher. I tried to get started several times during vacations or at conferences, when I had a break from the daily grind. But I never got past jotting down a page of ideas or a few paragraphs of prose before the never-ending academic demands of grant deadlines, article submissions, student supervision, and other responsibilities pulled me away. I was beginning to think the closest I might ever get to book authorship was working in my favorite spot at my neighborhood library, surrounded by amazing literature but with nothing to my name.

Then, my first daughter was born. Living in Canada, my partner and I had 1 year of paid parental leave between us. We decided my partner would take the first 6 months before returning to her job as a speech pathologist, and I would stay home with our daughter for the second half of the year. Suddenly, the work stopped. It was early in the pandemic, so we were stuck at home. My daughter did not mind if I took breaks from playing peek-a-boo, reading books, and snuggling so I could clean, cook, and do other chores. She also took consistent 2-hour naps. A few weeks into my leave, it occurred to me that I could use this nap time to get back to writing my book. That way, by the end of my leave, I might know whether book authorship was a real possibility or a dream I should leave behind, knowing at least I had tried.



"I worried whether ... I should be using this precious time for other tasks."

At first it was challenging to shift from caregiver to author and back again. I worried whether instead of indulging my yearning to write, I should be using this precious time for other tasks, such as reading more parenting books. But as the days passed, I felt more confident in my abilities as a father and writing became a welcome, if brief, change of pace. Quite often I was in the middle of jotting down an idea when that familiar wake-up cry reminded me I had other duties to attend to.

But I kept chipping away until, 2 months before I was due to return to work, I had a book proposal I was proud of. As I shared it with my most critical friends and received positive feedback, I grew determined to get the book published. With my partner's support, I plunged forward to the next big step: connecting with an agent and shopping around my proposal to publishers.

I was delighted to find an agent who worked with authors I admire and was interested in taking me on. As my parental leave came to a close and I returned to my day job, putting my book to the side, she sent my proposal to publishers. Then the rejections started coming in. The letters were kind, but they regularly explained that I was completely unknown outside my academic field and didn't have a "platform." I was accustomed to my academic articles being rejected, and I was always able to find a home for them somewhere. But as the rejections to my book proposal kept rolling in, I began to wonder whether this project had a chance. Finally, though, I got the call I was waiting for: My agent had found a publisher. I was ecstatic.

Now comes perhaps the biggest challenge: writing the book. I've realized the most important thing is to write every day. I couldn't do this without my partner's continued enthusiasm and encouragement, as well as taking on a little more of the household and child care responsibilities. I've also scaled back a bit on my academic work, including leaving some committees and not taking on new students to supervise. I feel lucky to be in a department, and at a stage in my career, where I have that freedom. I'm also grateful for the other advantages that have allowed me to get this far, including Canada's parental leave policy.

Even so, I've learned that no passion project can be done without trade-offs. I've learned that the odd quiet moments in a busy day can be a resource. And I've learned that I can only follow my dreams with the support of those closest to me. ■

Hasan Merali is a pediatric emergency medicine physician and associate professor of pediatrics at McMaster University. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER



Pushing the Boundaries of Knowledge

As AAAS's first multidisciplinary, open access journal, *Science Advances* publishes research that reflects the selectivity of high impact, innovative research you expect from the *Science* family of journals, published in an open access format to serve a vast and growing global audience. Check out the latest findings or learn how to submit your research: **science.org/journal/sciadv**

Science Advances
AAAS

GOLD OPEN ACCESS, DIGITAL, AND FREE TO ALL READERS

YOUR RESEARCH HAS ITS REWARDS.



Submit an essay by 3/15/23 to enter to win $25K.

The prestigious international *Science* & PINS Prize is awarded annually to outstanding early-career researchers in the emerging field of neuromodulation. Findings already show improved outcomes in several neurological disorders. To enter, just write a 1000-word essay about your research performed in the last three years. The Grand Prize winner will be presented a US$25,000 award, as well as have their essay published in *Science* magazine and on *Science* Online. A runner-up will have their essay published online. For additional inquiries, email SciencePINSPrize@aaas.org.

SCIENCE-PINS PRIZE FOR NEUROMODULATION

SCIENCE-PINS

SCIENCE.ORG/PINS