FDA, NIH fail to enforce reporting of clinical trials p. 240

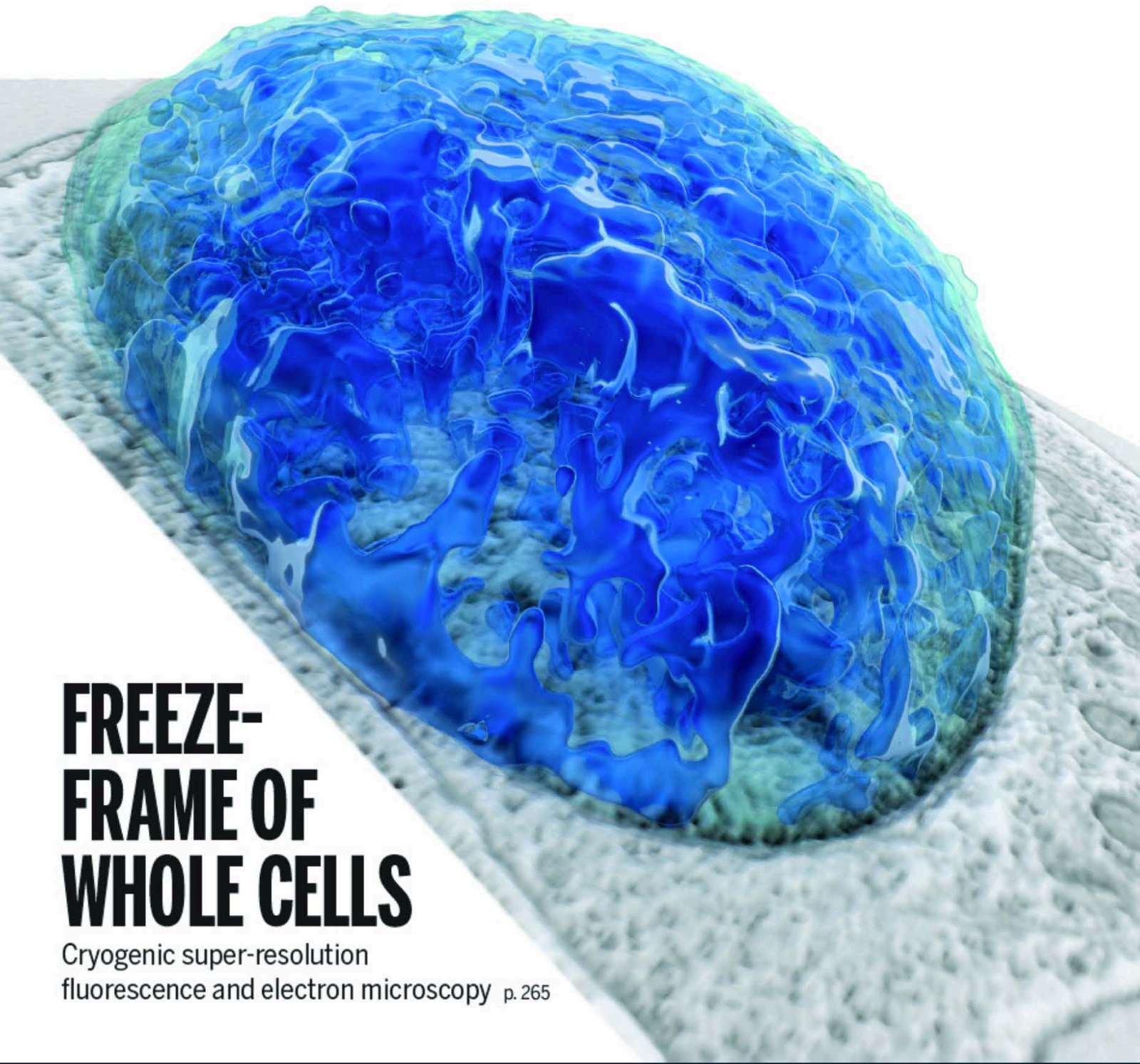
Microbial transmission in disease p. 250

Dynamic shape-changing wings p. 293

Science

$15
17 JANUARY 2020
sciencemag.org

AAAS



FREEZE-FRAME OF WHOLE CELLS

Cryogenic super-resolution fluorescence and electron microscopy p. 265

CONTENTS

17 JANUARY 2020 • VOLUME 367 • ISSUE 6475

255 Seeking our destiny in data

STOP THE PRESS

240

PHOTO: G. DAGLI ORTI/DE AGOSTINI PICTURE LIBRARY/GETTY IMAGES; ILLUSTRATION: DAVIDE BONAZZI/SALZMAN ART

NEWS

IN BRIEF

230 News at a glance

IN DEPTH

234 New SARS-like virus in China triggers alarm
Pneumonia outbreak in Wuhan appears to subside, but the virus could re-emerge
By J. Cohen and D. Normile

235 Electron-Ion Collider would lay bare the proton's innards
Department of Energy picks site for billion-dollar machine
By A. Cho

236 New front emerges in battle to build giant telescope in Hawaii
TMT opponents try to win over other astronomers in priority-setting decadal survey *By D. Clery*

238 In effort to save hemlocks, a rare glimpse of hope
Introduced predatory beetle eats deadly aphidlike pest, long-term study finds, but trees are still at risk
By G. Popkin

239 'Diversity statements' divide mathematicians
Hundreds choose sides both for and against requiring such statements in academic hiring
By M. Price

FEATURES

240 Transparency on trial
Many clinical trial results aren't posted publicly, as U.S. law requires—and a promised crackdown has fizzled *By C. Piller*

243 Gaming the system
By C. Piller
PODCAST

INSIGHTS

PERSPECTIVES

244 Eat and listen—how chewing and hearing evolved
Mammalian middle ear bones separated from the jaw of vertebrate ancestors *By J. A. Schultz*
REPORT p. 305

246 Water loss regulates cell and vesicle volume
Active control of vesicle water content ensures that cell size is maintained when extracellular fluid is taken up
By J. S. King and E. Smythe
REPORT p. 301

247 Enforcing T cell innocence
Inhibitory signaling on T cells maintains "naïvety" to self-antigen
By C. Brown and A. Y. Rudensky
RESEARCH ARTICLE p. 264

249 High-resolution dating of Paleozoic fossils
Refined ages of marine fossils clarify the timing of diversification and extinction events *By P. Wagner*
RESEARCH ARTICLE p. 272

250 Are noncommunicable diseases communicable?
Numerous noncommunicable diseases could have a transmissible microbial component
By B. B. Finlay et al.
PODCAST

POLICY FORUM

252 Mobile protected areas for biodiversity on the high seas
Protecting mobile marine species and habitats under climate change will require innovative and dynamic tools
By S. M. Maxwell et al.

BOOKS *ET AL.*

255 Patterns, statistics, and the stars
Highlights from the history of astrology reveal the origins of our quest to find meaning in data *By S. Vanden Broecke*

Published by AAAS

256 Knowledge in the Anthropocene
To save society, science should embrace ethical projects *By D. R. Coen*

LETTERS

257 Losing Mexican manatees
By G. Núñez-Nogueira and M. Uribe-López

257 Support Austria's glyphosate ban
By W. Peng et al.

258 Fund plant conservation to solve biodiversity crisis
By E. B. Roberson et al.

RESEARCH

IN BRIEF

260 From *Science* and other journals

REVIEW

263 Quantum materials
Quantum spin liquids
C. Broholm et al.
REVIEW SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.AAY0668

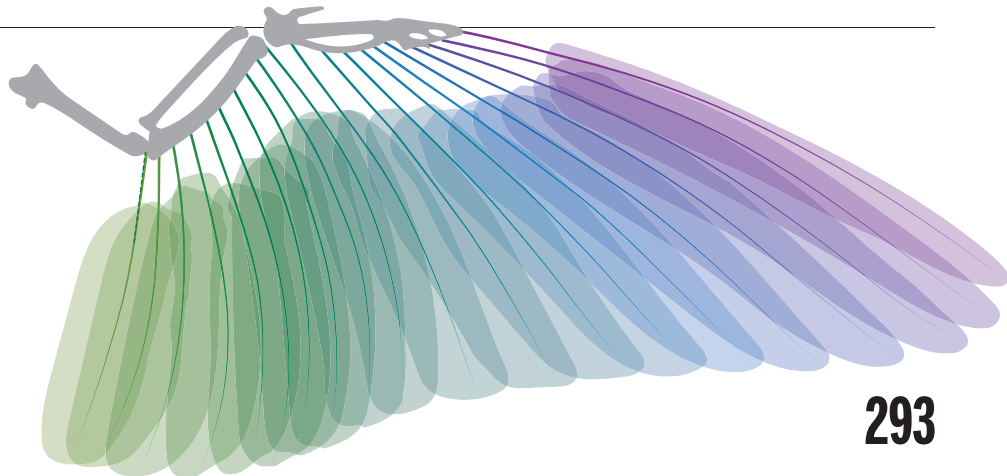
RESEARCH ARTICLES

264 T cells
VISTA is a checkpoint regulator for naïve T cell quiescence and peripheral tolerance
M. A. ElTanbouly et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.AAY0524
PERSPECTIVE p. 247

265 Advanced imaging
Correlative three-dimensional super-resolution and block-face electron microscopy of whole vitreously frozen cells *D. P. Hoffman* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.AAZ5357

266 Mass extinction
On impact and volcanism across the Cretaceous-Paleogene boundary
P. M. Hull et al.

272 Paleontology
A high-resolution summary of Cambrian to Early Triassic marine invertebrate biodiversity *J. Fan* et al.
PERSPECTIVE p. 249



293

REPORTS

278 Solar physics
Decay of the coronal magnetic field can release sufficient energy to power a solar flare *G. D. Fleishman* et al.

281 Organic chemistry
Nitromethane as a nitrogen donor in Schmidt-type formation of amides and nitriles *J. Liu* et al.

285 Quantum criticality
Singular charge fluctuations at a magnetic quantum critical point
L. Prochaska et al.

288 Optics
Subwavelength dielectric resonators for nonlinear nanophotonics
K. Koshelev et al.

293 Biomechanics
How flight feathers stick together to form a continuous morphing wing
L. Y. Matloff et al.
SCIENCE ROBOTICS RESEARCH ARTICLE BY E. CHANG *ET AL*. 10.1126/SCIROBOTICS.AAY1246

297 Martian atmosphere
Stormy water on Mars: The distribution and saturation of atmospheric water during the dusty season
A. A. Fedorova et al.

301 Cell biology
Lipid-gated monovalent ion fluxes regulate endocytic traffic and support immune surveillance *S. A. Freeman* et al.
PERSPECTIVE p. 246

305 Mammalian evolution
Integrated hearing and chewing modules decoupled in a Cretaceous stem therian mammal *F. Mao* et al.
PERSPECTIVE p. 244

309 Thermal conductivity
Phonon hydrodynamics and ultrahigh–room-temperature thermal conductivity in thin graphite *Y. Machida* et al.

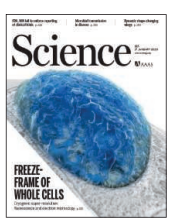
313 Organometallics
Fluorination of arylboronic esters enabled by bismuth redox catalysis
O. Planas et al.

DEPARTMENTS

229 Editorial
Do bans help modern public health?
By Lawrence O. Gostin

334 Working Life
Getting personal *By Erin Gibson*

ON THE COVER



Translucently colored nucleus of a mammalian cerebellar granule neuron. This 3D rendering illustrates the specific heterochromatin subdomain type defined by the overlap of histones imaged by electron microscopy (EM) and cryogenic super-resolution fluorescence microscopy. A thin translucent shell surrounding the nucleus represents the nuclear membrane, and a thin slab of 3D-rendered EM data (gray) slices through the nucleus. The circles at lower right are cross-sectional views of mitochondria. See page 265.
Image: D. Hoffman, G. Shtengel, E. Betzig, H. Hess, Janelia Research Campus; K. Campbell, D. Stabley, D. Solecki, St. Jude Children's Research Hospital

New Products 318
AAAS Meeting Program 319
Science Careers 329

ILLUSTRATION: L. Y. MATLOFF *ET AL.*

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 © 2020 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): $2148; Foreign postage extra: Air assist delivery: $98. 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.sciencemag.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.

Published by AAAS

EDITORIAL

Do bans help modern public health?

A century ago, the 18th Amendment to the United States Constitution went into effect, banning the "manufacture, sale, or transportation of intoxicating liquors." Fourteen years after its ratification, the 18th Amendment was repealed by the 21st Amendment. What did Prohibition teach us about banning hazardous products like alcohol, tobacco, or e-cigarettes?

The 18th Amendment was a failed "noble experiment," with unforeseen harms, including a thriving black market, organized crime, and sporadic enforcement. Eventually, illicit sale of liquor became easily affordable. Pervasive flouting of Prohibition undermined the rule of law.

A prohibition on hazardous activities is a blunt tool because products often have both public health risks and benefits. Using illicit drugs is addictive and harmful, but needle exchanges can reduce harms. E-cigarettes can cause acute and longer-term hazards, but they can help cigarette smokers to quit. If government bans a product, it cannot tax it, thus forgoing vital revenues. Lawful marijuana sales in the United States, for example, have financed public services, such as education.

Alcohol is poured down sewers during prohibition days in the United States.

There are no easy answers, but strict regulation of unsafe products is a more flexible tool to decrease behavioral risks, while avoiding social harms (a black market or discriminatory enforcement). Regulations are often more politically viable than bans, which raise concerns about paternalism and the "nanny state."

Tobacco control offers a paradigmatic case of effective rules. A suite of measures, including taxes, age limits for purchasing, marketing restrictions, graphic warnings, and public smoking curbs, has greatly reduced smoking rates. The World Health Organization's Framework Convention on Tobacco Control codified this regulatory system globally. Similar public health benefits could be achieved by controlling other unhealthy products, including alcoholic beverages, "junk" foods, and sugar-sweetened beverages. Taxes, for example, have been shown to reduce consumption of the latter. Gradually reducing sodium in packaged foods could lower hypertension rates.

Government sometimes criminalizes activities without any evidence of harm. Bans on needle exchange programs, for example, are counterproductive. Research shows that exchanges do not encourage drug use but, rather, reduce the sharing of contaminated drug injection equipment.

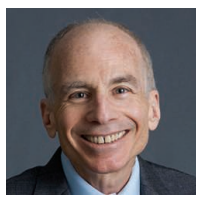
The harder cases entail products that have a dual use, causing harm to some consumers, while safeguarding others. Debate swirls around banning e-cigarettes, which can cause lung damage and nicotine poisoning. The benefits and harms of vaping are not fully understood, but evidence suggests that vaping could be a gateway to tobacco use; it also could serve as a harm reduction strategy for tobacco smokers. Prohibiting vaping would cause a public backlash and extinguish any benefit from harm reduction. A suite of regulations would be more nuanced, including taxes, age restrictions for purchasing, youth marketing curbs, outlawing all flavors, and even requiring a physician's prescription to purchase e-cigarettes.

Marijuana laws stir public controversy, but there is also incomplete evidence regarding the health benefits and harms. Government strategies are inconsistent: U.S. federal law bans all marijuana use, whereas many states allow marijuana for personal use or require a medical prescription. The majority of drug arrests in the United States are for marijuana, and mostly for simple possession. Discriminatory enforcement has led to disproportionate incarceration rates among African Americans.

Bans have another downside. Researchers can assess the effectiveness of regulations, but once government prohibits an activity, it becomes hard to evaluate. Evidence of effectiveness enables government to alter policies to safeguard the public's health.

Prohibition taught society to be cautious about bans. It is deceptively simple to criminalize a hazardous activity. But bans can create unforeseen social and political risks. The public does not support a government that tells individuals what they can or cannot do for their health. Yet government's greatest responsibility is to safeguard the public's health. It can do that through a well-regulated society—that is, with evidence-based interventions to "nudge" the public to adopt healthier and safer behaviors.

– Lawrence O. Gostin



Lawrence O. Gostin is the O'Neill Chair in Global Health Law and a professor of medicine at Georgetown University in Washington, DC, USA. He is director of the O'Neill Institute for National and Global Health Law at Georgetown Law Center, Washington, DC, USA, and director of the World Health Organization Collaborating Center on National and Global Health Law. gostin@georgetown.edu

PHOTOS (TOP TO BOTTOM): GEORGETOWN UNIVERSITY LAW CENTER; SHAWSHOTS/ALAMY STOCK PHOTO

10.1126/science.aba8857

Published by AAAS

NEWS

> “This policy is as clear as mud.”
>
> **Immunologist Scott Kitchen**, in *The Washington Post*, about the Trump administration’s 2019 guidelines on U.S. funding for research using fetal tissue from elective abortions. An ethics panel to review new research proposals and renewals has not been appointed.

IN BRIEF

Edited by
Jeffrey Brainard

Members of Montreal’s Iranian community attend a vigil for the airplane crash victims.

PHOTO: ANDREJ IVANOV/THE CANADIAN PRESS/AP IMAGES

DISASTERS

Researchers die in Iranian missile strike

Several dozen researchers and graduate students were killed on 8 January when Iran shot down a passenger jet shortly after it took off from Tehran, killing all 167 passengers and nine crew members on board. The Iranian government first denied responsibility but later acknowledged accidentally targeting the plane. Many victims held dual citizenship in Canada and Iran and were working or studying at Canadian universities; they were returning after visiting Iran during a holiday break. They included two engineering professors at the University of Alberta, Edmonton: Mojgan Daneshmand, an internationally recognized expert in radio frequency microsystems, and her husband, Pedram Mousavi. In all, 10 Canadian universities have reported losing students and faculty members in the crash. A full list of all victims was not available when *Science* went to press.

Second hottest year was 2019

CLIMATE | This past year was the planet’s second hottest in modern history, after 2016, the European Union’s Copernicus Climate Change Service announced last week. Driven by humanity’s ever-rising emissions of greenhouse gases, the 2019 surface temperatures capped a stretch of 5 years that have been the warmest on record, averaging 1.1°C to 1.2°C above preindustrial levels. The year saw near-record summer melt in Greenland and ice-free spots in the Arctic that lingered into the fall. For some regions, including Europe and Australia—now battling devastating bush fires—2019 was their warmest on record. What is more, heat in the world’s oceans set a new record high in 2019, according to a separate analysis published on 13 January in *Advances in Atmospheric Sciences*. The seas absorb 90% of the heat trapped by rising greenhouse gases, and ocean temperatures have shown a steadier upward climb than surface ones.

Indonesian leader’s cure criticized

BIOMEDICINE | Indonesian Minister of Health Terawan Agus Putranto alarmed the country’s medical community last month by suggesting that hospitals nationwide start to use a controversial treatment for stroke

Published by AAAS

that he pioneered. A radiologist by training, Terawan claims he has successfully treated thousands of patients since 2012 with intra-arterial heparin flushing (IAHF), which uses a blood thinner to reduce clots that may have formed during strokes. Indonesian neurologists have argued that there's no evidence that the therapy works and pointed out that it carries a small risk of complications and even death. Terawan's critics say his advocacy for the therapy disqualifies him from setting health policy for the world's fourth most populous nation. The Ministry of Health did not respond to a request for details about IAHF's proposed rollout.

New virus? Yada yada

MICROBIOLOGY | The researchers had to admit their discovery wasn't all that exciting. The team at AgriBio, the Centre for AgriBioscience in Melbourne, Australia, had found a new alphavirus in RNA from large numbers of mosquitoes collected in Australia's Victoria state. Although some viruses in that group cause human epidemics, the new one infects only mosquitoes—and it isn't the first insect-only alphavirus. Although co-discoverer Jana Batovska says the new virus might aid knowledge of virus evolution and vaccine production, she and colleagues chose a prosaic name for it in their short paper published 9 January in *Microbiology Resource Announcements*: Yada Yada virus. The catch phrase was made famous in a 1997 episode of *Seinfeld* to tar boring, empty talk.

POLAR SCIENCE

Brazil opens new Antarctic research center

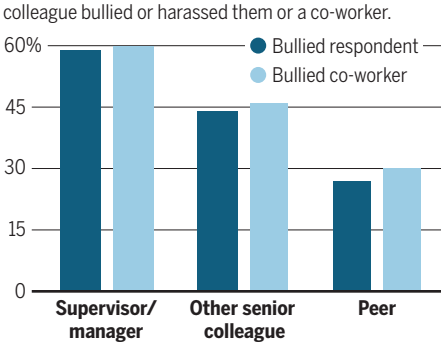
Brazil planned to dedicate a new scientific outpost in Antarctica this week, 8 years after a fire destroyed its original base there. The new, $100 million Comandante Ferraz Antarctic Station (above) is nearly twice the size of the old one and stands out for its sleek architectural design. It can house up to 64 people, including scientists and military personnel; 17 laboratories will support research in a range of fields, including environmental microbiology, human physiology, paleontology, and climate change. Since the fire, Brazilian scientists have continued to do research in a series of container modules at the site on the Keller Peninsula of King George Island. But many scientists worry that the Brazilian government—whose spending for research has plummeted in recent years—may not provide enough funds for the new facility to use its full potential. Currently allocated funds for Brazil's Antarctic research program are set to expire in 2022.

Survey finds bullying pervasive

WORKPLACE | Many researchers are proud and passionate about their work, but workplace problems are sapping research quality and leaving them feeling stressed and lonely, says a report this week from the Wellcome Trust. The findings come from what the London-based biomedical charity describes as the largest study of views about workplace culture among researchers at different career stages in multiple disciplines. The survey's more than 4200 respondents included some working outside the biomedical sciences, including in earth science and the humanities; 76% worked in the United Kingdom. Four-fifths believe that high levels of competition for grants and jobs are creating unkind or aggressive working conditions; many reported being bullied by colleagues. The report suggests possible methods for improvement, such as institutional ombudsmen to collect and consider complaints.

▲ Who the bullies are

More than four in 10 survey respondents reported that a colleague bullied or harassed them or a co-worker.



CREDITS: (PHOTO) ESTÚDIO 41; (GRAPHIC) J. BRAINARD/*SCIENCE*; (DATA) WELLCOME TRUST

MIT details Epstein donations

#METOO | Jeffrey Epstein donated $750,000 to the Massachusetts Institute of Technology (MIT) after his 2008 criminal conviction for procuring a minor for prostitution, and three MIT vice presidents knew about and approved most of the gifts, according to a report by an outside law firm. The report, issued on 10 January, says Epstein visited the campus at least nine times from 2013 to 2017 and met once with students, without the knowledge of senior administrators. The postconviction donations "were driven either by former Media Lab Director Joi Ito or by [mechanical engineering] Professor [Seth] Lloyd," the report finds. MIT President Rafael Reif announced on 10 January that Lloyd, who received $225,000 in research support from Epstein and a personal gift of $60,000, has been put on paid administrative leave. Ito resigned in September 2019. Epstein died of an apparent suicide in federal custody in New York City in August 2019 while awaiting trial on charges of sex trafficking minors.

Florida probes foreign influences

RESEARCH COLLABORATION | Florida's legislature is diving into the contentious issue of foreign influences on U.S. research. The move comes after the Moffitt Cancer Center in Tampa last month fired six scientists, including CEO Alan List, for failing to disclose ties to Chinese institutions. A special committee will begin hearings next

Published by AAAS

THREE QS

Science boosts sci-fi

The fourth season of *The Expanse*, a hard-boiled TV space drama known for its real-world space physics, debuted last month. Executive Producer Naren Shankar chatted with *Science* about using his doctorate in applied physics and electrical engineering to help get the science right. (A longer version of this interview is available at https://scim.ag/TheExpanse.)

Q: What did your science education bring to your TV work?
A: One of the most valuable things I took away from school is peer review. You write a paper, sit down with your colleagues, and then you pare it down. That is really the process of writing a script. I did a lot of science fiction in the early stage of my career, and then cop and crime shows. *CSI*—I ran that show for many years. It had a lot of scientific method in it. The idea of the logical path to do a criminal investigation, evaluating evidence.

Q: What's realistic about the physics of spacecraft in the show?
A: The fact that you don't have weight unless your ship is accelerating, the fact that communication in space is not instantaneous. We use that for drama. At the end of one episode, missiles are heading off to hit Mars. In the very next episode, people on Earth are realizing that [the missiles struck], like, 25 minutes ago. You see conservation of momentum, conservation of angular momentum: all the things that would actually occur in space. You don't see control surfaces and aerodynamic flight, because they're all moving in a vacuum. You see realistic objects changing orientation with thrusters. Personally, I'm quite tired of seeing spaceships fly around like fighter planes in the Pacific in World War II.

Q: Does the show's realistic science fire up the physicist in you?
A: It's actually one of the things that attracted me to the project. It embraced all of the things that most science fiction shows run away from. It requires appreciation of the reality of what's going on. And so it's a fairly high bar, I think. But it's certainly not inaccessible.

S SCIENCEMAG.ORG/NEWS
Read more news from *Science* online.

A massive assembly will hold the telescope's 8.4-meter-wide mirror.

ASTRONOMY

A telescope with three names

A U.S.-funded telescope in Chile that will scan the entire sky every 3 days for sudden changes has experienced a sudden change of its own—acquiring three new names. The $1.4 billion Large Synoptic Survey Telescope (LSST), as it has been known since the National Science Foundation (NSF) approved it in 2012, is to begin operations in 2022. Such facilities are often renamed on completion, and NSF polled the project team for ideas. Many wanted to name it after Vera Rubin, the pioneering astronomer who first identified the effects of dark matter and who died in 2016. Some wanted to retain the name LSST. Then, Congress passed a law last month naming it the Vera Rubin Survey Telescope, but NSF also wanted to mark the $20 million that philanthropist Charles Simonyi gave the project early on to start work on its unique 8.4-meter, wide-viewing mirror. So—take a deep breath—the project's initial 10-year mapping mission is now called the Legacy Survey of Space and Time, and it will be carried out by the Charles Simonyi Survey Telescope at the NSF Vera C. Rubin Observatory. The announcement, made last week at the annual meeting of the American Astronomical Society in Honolulu, marks the first time a national observatory has been named after a woman.

week into such collaborations by scientists at any research institution receiving state funding, making Florida the first state to raise the same type of questions that federal research agencies have been asking their grantees. Moffitt officials say the fired scientists had secretly accepted money for participating in Chinese talent recruitment programs. Although Moffitt says it found no evidence that the collaborations led to the illicit transfer of intellectual property, state Representative Chris Sprowls (R), who is leading the probe, worries "the intent is clearly there."

U.S. eyes nixing climate analysis

REGULATION | Under a major change to environmental policy proposed last week by the Trump administration, U.S. federal agencies reviewing proposed pipelines and other large construction projects would not have to consider their impact on climate. The National Environmental Policy Act (NEPA), enacted in 1970, requires officials to evaluate the potential environmental damage, such as water pollution or harm to endangered species, from activities that require government permits. In 2016, then-President Barack Obama's administration added climate change to the list of impacts, including the effects on climate of emissions stemming from development of oil, gas, and coal reserves. The policy change is intended to speed up the NEPA reviews and limit them to 2 years. Public comments are due by 10 March.

PHOTO: LSST PROJECT/NSF/AURA

Published by AAAS

On New Year's Day, Wuhan health authorities closed a live animal market linked to the mysterious outbreak.

PHOTO: NOEL CELIS/AFP/GETTY IMAGES

INFECTIOUS DISEASES

New SARS-like virus in China triggers alarm

Pneumonia outbreak in Wuhan appears to subside, but the virus could re-emerge

By **Jon Cohen** *and* **Dennis Normile**

Had the nightmare returned? That's the question many were asking in the first 10 days of this year, after a new form of pneumonia emerged in Wuhan, a megacity in central China. The outbreak revived memories of severe acute respiratory syndrome (SARS), the disease that emerged in China in 2002 and sickened 8098 people in 37 countries before it was quashed in the summer of 2003. Like SARS, the Wuhan pneumonia cases were linked to a market selling myriad species of live animals, and they appear to be caused by a new member of the coronavirus family closely related to the SARS virus. And once again, China appeared to be less than forthcoming with information.

Today, global health experts are breathing a little easier. As *Science* went to press, only one of 42 people known to be infected had died: a 61-year-old man already suffering from abdominal tumors and chronic liver disease. (SARS had a 9.6% mortality rate.) No evidence suggests the virus easily passes between humans, which can turn a local problem into a global crisis. And Chinese researchers have now shared the sequence of six genomes of the as-yet-unnamed virus with the world, which scientists elsewhere have used to quickly develop and publish a diagnostic test. Ralph Baric, a coronavirus researcher at the University of North Carolina, Chapel Hill, is already trying to synthesize live virus from the data so that he can study it in animals.

Still, many questions remain. Researchers have not identified the animal species at the marketplace that harbored the virus. When it emerged and the true number of people infected remain a mystery. Meanwhile, a case in Thailand, reported on 13 January—in a tourist who flew from Wuhan to Bangkok—led World Health Organization (WHO) Director-General Tedros Adhanom Ghebreyesus to consult experts on outbreak responses. The patient had not visited the Wuhan market at the center of the outbreak but had been to other animal markets, suggesting the virus has spread within Wuhan, the *South China Morning Post* reported on 14 January.

The first known patient developed symptoms—which can include difficulty breathing and fever—on 8 December 2019. Officials closed the seafood market on New Year's Day, and no new patients have been identified in Wuhan since 3 January. The virus was not found in 763 close contacts of those infected, or in health care workers, who often fall ill during outbreaks of viruses that can transmit between humans. "It is a limited outbreak," says Xu Jianguo, who runs an infectious disease laboratory at the Chinese Center for Disease Control and Prevention and heads an evaluation committee that's advising the Chinese government. "If no new patients appear in the next week, it might be over."

WHO said in a 12 January statement that it was "reassured of the quality of the ongoing investigations and the response measures implemented in Wuhan, and the commitment to share information regularly."

But others criticized the way early information came out. News that researchers had discovered a novel coronavirus came in an 8 January story in *The Wall Street Journal*; Xu confirmed the finding on a state-run TV station several hours later. "It's not a good situation when *The Wall Street Journal* [reports] a SARS-like coronavirus before the Chinese government announces it," Baric says. On 10 January, Jeremy Farrar, an infectious disease specialist who heads the London-based Wellcome Trust, tweeted his worry about rumors that the Chinese government did not share "critical public health information" because Chinese researchers wanted to ensure publication of their findings in high-profile journals first.

Less than 12 hours later, however, evolutionary biologist Edward Holmes of the University of Sydney published an "initial"

Published by AAAS

sequence of the new coronavirus on virological.org, on behalf of a consortium led by Zhang Yong-Zhen of Fudan University in Shanghai. The next day, three groups working under China's National Health Commission posted another five sequences of the virus, gathered from different patients, on GISAID, a database primarily used for sharing data on influenza viruses.

The six sequences differ little from each other, which evolutionary biologist Andrew Rambaut of the University of Edinburgh says is "consistent with a point source"—meaning they likely came from the same batch of infected animals at the Huanan Seafood Wholesale Market, which also sells birds, snakes, and rabbit meat. (No coronaviruses have ever been found in fish.) But the fact that cases surfaced over the course of 1 month suggests the source was more than one group of animals at one location, Farrar says: "It makes me worry that whatever the exposure was to, it went on for quite a long time." Virologist Guan Yi of Hong Kong University agrees that the Wuhan outbreak was caused by multiple jumps from animal to human hosts "separately and independently."

Whatever species spread the virus at the market may have picked it up from some natural reservoir. Many coronaviruses occur naturally in bats, and the new virus is closest to four bat viruses that have surface proteins capable of infecting human cells. Still, Rambaut cautions there may well be another natural host. "It's quite similar to a bat virus in parts of its genome, but not so much in other parts," he says.

Farrar notes that most confirmed cases to date were mild, which means that even before health officials recognized the outbreak, the virus may have infected many other people who never sought medical care. That makes it premature to conclude the pathogen doesn't spread from human to human, he says. Nurses and doctors, too, may have been infected without anyone noticing, he adds: "With a coronavirus, I'd be very surprised if there wasn't some limited human-to-human transmission." So far, cases have been confirmed by detecting nucleic acid from the virus, which disappears after patients recover. Now that the virus has been isolated, researchers can also develop antibody tests that pick up signs of past infection.

Limited as the outbreak appears to date, Farrar and others still worry that travel of hundreds of millions of people for the Lunar New Year celebration on 25 January could spread the virus from Wuhan, a major transportation hub, to other cities. "With people, food and animals move," says Farrar, who suspects that this outbreak "is not going away anytime soon." ■

NUCLEAR PHYSICS

Electron-Ion Collider would lay bare the proton's innards

Department of Energy picks site for billion-dollar machine

By **Adrian Cho**

The United States has taken a key step toward building its first new particle collider in decades. Last week, the U.S. Department of Energy (DOE) announced that the Electron-Ion Collider (EIC) will be built at Brookhaven National Laboratory in Upton, New York. The machine would enable nuclear physicists to probe the mysterious structure of the proton and how its mass and spin emerge from a teeming sea of even smaller subatomic particles inside it.

"The U.S. has been at the front end in nuclear physics since the end of the Second World War and this machine will enable the U.S. to stay at the front end for decades to come," said Paul Dabbar, DOE's undersecretary for science, in a telephone press conference announcing the site selection for the machine, which will cost between $1.6 billion and $2.6 billion and could begin to run by 2030. DOE's Thomas Jefferson National Accelerator Facility in Newport News, Virginia, had also vied to host it.

For decades, physicists have fired electrons into protons and atomic nuclei. In the early 1970s, these experiments showed that each proton (and neutron) consists of three less massive quarks, which bind to one another by exchanging quantum particles called gluons.

However, quantum uncertainty causes the proton's interior to roil with countless gluons, quarks, and antimatter antiquarks that flit in and out of existence too quickly to be directly observed. Many of the proton's properties—including its mass and spin—emerge from that sea of "virtual" particles in ways that theorists don't understand, says Gordon Baym, a nuclear theorist at the University of Illinois, Urbana-Champaign, who led a 2018 study by the National Academies of Sciences, Engineering, and Medicine that called for an EIC (*Science*, 27 July 2018, p. 317). "What is [the gluons'] distribution in space? What is their distribution in momentum?" he says. "We don't know much about that."

It's not for lack of trying. Since 1994, the Continuous Electron Beam Accelerator Facility (CEBAF, pronounced "see-baff") at Jefferson lab has fired electrons into targets rich in protons and neutrons. But CEBAF can only probe the more energetic of the proton's virtual quarks and gluons, those that individually carry more than about 20% of the proton's total momentum.

With its more intense and energetic beams, the EIC should see the more numerous quarks and gluons that carry as little as 1/100,000 of the proton's momentum. That throng of gluons should crowd together so much that their identities as individual particles blur as they form a new state of matter called a color-glass condensate, says Peter Braun-Munzinger, a high-energy and nuclear physicist at the GSI Helmholtz Center for Heavy Ion Research in Darmstadt, Germany.

The Electron-Ion Collider would add an electron beam to the Relativistic Heavy Ion Collider ring in New York.

PHOTO: BROOKHAVEN NATIONAL LABORATORY/FLICKR CC BY-NC-ND

Published by AAAS

Previous experiments have hinted at the existence of the condensate, but the EIC could clinch the case. The state's subtle signals should emerge more strongly in collisions with heavier nuclei, Braun-Munzinger says. That's why in addition to accelerating protons, the collider also needs to be able to accelerate nuclei, or ions.

Dabbar declined to say why DOE chose Brookhaven over Jefferson lab, but the Virginia facility was playing David to Brookhaven's Goliath. Brookhaven covers more than 2100 hectares and has 2500 employees. Jefferson lab has a staff of fewer than 700 and occupies 68 hectares. Its researchers had hoped to build the EIC by adding an ion accelerator to CEBAF. But compared with an electron accelerator, an ion accelerator is larger and more expensive, and Jefferson's small campus would have limited the ion accelerator's size and energy.

Brookhaven, in contrast, already has a massive ion collider, housed in a 3.8-kilometer tunnel. Since 2000, the Relativistic Heavy Ion Collider (RHIC, pronounced "rick") has smashed together nuclei such as gold and copper to re-create in fleeting puffs the ultrahot plasma of quarks and gluons that filled the newborn universe. To build the EIC, researchers at the Long Island lab plan to add an electron accelerator in the same tunnel as RHIC.

The EIC won't be easy to build. To reach the desired high collision rate, the machine will have to concentrate its proton or ion beams in unprecedented ways. To probe the origins of the proton's spin, it will have to polarize its colliding beams to a degree never before achieved, says Lia Merminga, an accelerator physicist at Fermi National Accelerator Laboratory. "This requires techniques that are beyond the state of the art," she says.

At the moment, the EIC's only competition is a plan by researchers in China to build an electron-ion collider as part of a new physics lab in Huizhou. But that machine would run at a much lower energy than the EIC and focus more on quarks than gluons, making the two machines complementary, says Xurong Chen, a physicist at the Chinese Academy of Sciences's Institute of Modern Physics.

To win DOE's approval for construction, Brookhaven researchers must develop a detailed design and cost estimate. They plan to complete that work by 2024, when RHIC shuts down, says Brookhaven Director Doon Gibbs. With DOE science funding up 31% since 2016, Dabbar is optimistic that the agency will see the EIC through. "We've been able to start on every major project that's been on the books for years," he says. ■

ASTRONOMY

New front emerges in battle to build giant telescope in Hawaii

TMT opponents try to win over other astronomers in priority-setting decadal survey

By **Daniel Clery**, *in Honolulu and on the slopes of Mauna Kea in Hawaii*

On the rain-lashed road leading to the peak of Mauna Kea on Hawaii's Big Island, the sounds of singing and drumming can be heard above a roaring wind. Under a canopy, dozens of people celebrate their connection to the mountain with Native Hawaiian songs, chants, and dances, as they have been doing three times daily for the past 7 months. Many of the tents scattered across this lava field are unoccupied during this winter lull, but the current residents take good care of the site, which includes tents for food, medical care, and classes. They seem happy to wait as long as it takes for astronomers to give up on their plan to build a giant telescope on their sacred mountaintop.

> **"If we put out a call, 1000 people would be here in an hour."**
> **Uncle Sparky**, Mauna Kea protector

Like medieval armies calling a halt to war in winter, the two sides contesting the soul of Mauna Kea agreed last month to a 2-month truce: The consortium that wants to build the Thirty Meter Telescope (TMT) here will not attempt to start construction, and the camp residents—who see themselves as the mountain's protectors, or *kia'i*—moved their tents from the road, allowing unimpeded access up and down the mountain. "A de-escalation of emotions is a good thing," says Gregory Chun, a psychologist at the University of Hawaii (UH), Hilo, who is advising the university leadership on the management of Mauna Kea. "I don't know what will come of it."

Behind the scenes, TMT board members and funders are meeting with Native Hawaiian leaders—some of whom support the project—but this has yet to yield any solutions. Few can predict what might happen when the truce ends in late February, but no one doubts that *kia'i* would return in force if the TMT consortium tries to start construction, says a longtime *kia'i* known as Uncle Sparky. "If we put out a call, 1000 people would be here in an hour."

And now, the $1.4 billion TMT, backed by six nations and wealthy universities, is facing another front in its battle to be one of three new giant telescopes—and the only one in the Northern Hemisphere. TMT opponents—including some astronomers—are seeking to win over the wider astronomy community to stop its construction.

In a press conference outside the annual meeting of the American Astronomical Society (AAS) in Honolulu last week, a group of Hawaiian researchers announced the submission of eight white papers to the decadal survey in astronomy, known as Astro2020, a priority-setting exercise that influences U.S. funding agencies. They want Astro2020 to ensure that no federal money is used to build on state land without the consent of local Indigenous people. And federal money is just what the TMT needs. In a joint pitch to Astro2020 with one of its two rivals, the U.S.-led Giant Magellan Telescope in Chile, the projects are asking for money that would fill budget holes and provide observing time for U.S. astronomers.

The white papers describe the cultural significance of Mauna Kea and the negative impact of the observatories on Indigenous people and their ability to carry out cultural practices on the mountain. One paper charts the decades of resistance—both legal and direct action—to telescopes on Mauna Kea. The mountain and other land throughout the islands that once belonged to the Hawaiian monarchy was seized after its overthrow in 1893 and became state land when Hawaii was granted U.S. statehood in 1959. "It's a land rights issue, and a power and decision-making issue," says co-author Shelley Muneoka, a sociologist at the Hawaiian-Environmental Alliance.

In July 2019, after surmounting legal challenges, TMT officials were finally ready to begin construction. When police arrived intending to ensure the safe passage of construction traffic, *kia'i* had already chained themselves to a cattle grid on the access road. Over a few chaotic days, police arrested 38 people, many of them community

Published by AAAS

The TMT (artist's concept) would be the biggest telescope on Mauna Kea, a peak many Native Hawaiians hold sacred.

IMAGE: TMT INTERNATIONAL OBSERVATORY

elders, or *kupuna*—which only hardened the opposition on the mountainside and attracted wider interest, including from actors and celebrities who joined the protests. "Arresting elders really changed people's minds," says TMT opponent Rosie Alegado, an oceanographer at UH Manoa.

TMT officials have said they can't condone the use of force to get the telescope built. Canada, one TMT partner, is particularly concerned because of its own checkered history with Indigenous people, Canadian astronomers say. Construction will take most of a decade and require more than 2000 truck trips up the mountain. If protesters constantly harry that effort, an accident is inevitable.

The 3500 astronomers who arrived last week for the AAS meeting found themselves in an awkward situation: wanting to promote the best in astronomy but sensitive to the strong feelings in the islands about building on its mountaintops. To quote one AAS delegate, the TMT was the "30-meter elephant in the room."

AAS attendees did not see thousands of TMT protesters, like the crowd that greeted the International Astronomical Union when it came to the same venue in 2015. Instead, *kia'i* were invited to join debates and speak in sessions. "The impact on our Mauna is an impact on our very soul," *kia'i* leader Noe Noe Wong-Wilson told an Astro2020 session on diversity at the meeting. "We ask TMT not to build. It's not about the science. It is sacred. It belongs to the gods."

Native Hawaiian and nonastronomer supporters of the TMT also made their voices heard. For them, it is about well-paid technical jobs in a state with high living costs and scant opportunities. Tourism, a major source of jobs, has in recent years been buffeted by rises in fuel costs, financial crises, and volcanic eruptions. "Science will always be stable," says Malia Martin, founder of Imua TMT (*imua* means "forward"). She sees telescopes on Mauna Kea as "an opportunity for Indigenous people to command a global industry." The opponents "protested and lost, and they need to accept that," says Imua TMT's Amber Imai-Hong, an aerospace engineer at UH's Hawaii Space Flight Laboratory.

Many astronomers still dearly want the TMT built, and take issue with the misinformation spread on social media by some opponents, including claims that it will pollute the island's groundwater, that it will be nuclear powered, and that it is a laser weapon built by China. But many would not comment publicly. TMT officials kept a low profile at the meeting, only holding off-the-record meetings with journalists.

Robert McLaren, director of UH's Institute for Astronomy, is concerned that the opposition to astronomy could threaten the dozen other large observatories already on Mauna Kea. The university manages the mountaintop under a lease from the state that must be renewed before the end of 2033. The institute has been working on its application for 5 years, carrying out an environmental impact survey and revising its management plans, McLaren says. He hasn't yet sensed that opponents are targeting the lease renewal, but says for some, at least, "their ultimate goal is to have nothing there."

Another worry is financial. The observatories pay just $1 a year for their subleases, but they share day-to-day costs such as road maintenance, snow removal, and operation of the visitor center. Nevertheless, UH chips in $2 million per year—an unsustainable load, McLaren says. The observatories expect to pay higher rent after 2033, but the arrival of the TMT could strain their finances. To appease opponents, Governor David Ige (D) in 2015 called on UH to decommission older and less productive telescopes. The university is in the process of identifying up to five for closure, which would mean fewer to share the burden of operating the site. "How many are going to be at the party to pay the bills?" McLaren asks.

For the TMT itself, it is a waiting game. The project funders—the California Institute of Technology, the University of California, Japan, China, India, and Canada—have to decide how long they can wait before considering the backup site of La Palma in Spain's Canary Islands. For astronomy, the site is inferior, but it is less politically fraught. Every passing month adds costs. Glass blanks for more than 60% of the telescope's 574 1.4-meter-wide mirror segments sits in warehouses. Polishing has started in the United States and Japan and will soon begin in India and China. Designs for the telescope structure and its enclosure are ready, but contractors have been put on hold.

John Evans, a *kia'i* at the Mauna Kea camp, is happy to keep waiting. The wiry 68-year-old with a long white beard is a recent convert to Hawaiian dancing, and wants to keep at it. "I wish them well," he says, of the TMT. "Just somewhere else." ■

Published by AAAS

An introduced beetle that eats the eggs of the hemlock woolly adelgid is showing promise.

BIOCONTROL

In effort to save hemlocks, a rare glimpse of hope

Introduced predatory beetle eats deadly aphidlike pest, long-term study finds, but trees are still at risk

By **Gabriel Popkin**

A potential ally for one of North America's most embattled trees has passed its first big test. A tiny predatory beetle that researchers have been rearing and releasing into forests appears to be doing damage to an aphidlike pest that poses a deadly threat to ecologically important eastern hemlock trees, a 5-year study has found.

The result marks a rare success for forest scientists aiming to use introduced insects to battle pests, a strategy called biocontrol. Researchers caution that hemlocks are far from safe, however, because it is unlikely the beetle alone can defeat the pest. But the news "gives some cause for encouragement," says Aaron Ellison, an ecologist at the Harvard Forest, who is not involved in the work.

Sometimes called the redwood of the east, eastern hemlock (*Tsuga canadensis*) is an evergreen giant inhabiting forests in eastern North America from Alabama to southern Canada. Its dense shade cools mountain streams, and its leaves and branches host hundreds of insects and other arthropods. It can form almost pure stands in the forests of the northern United States and Canada, making it a foundational species.

Since the 1980s, however, hemlocks have come under an ever-widening assault from the hemlock woolly adelgid, a tiny insect native to Japan that sucks sugars from hemlock needles, killing trees (*Science*, 21 August 2015, p. 802). The adelgid has left ghost forests throughout the Appalachian Mountains and southern New England. It recently showed up in Michigan, putting hundreds of millions of trees in the upper Midwest at risk.

To save the hemlock, many scientists have placed their bets on biocontrol: introducing one or more species to reduce adelgid populations. In the 1990s, investigators in British Columbia in Canada found that a black beetle named *Laricobius nigrinus*, about the size of a grain of rice, munches on adelgid eggs, larvae, and nymphs. Scientists have since reared and released more than 400,000 of the beetles at sites throughout the eastern hemlock's range and watched some populations take hold.

To see whether the beetles were actually harming the adelgid, a team led by entomologist Scott Salom of Virginia Polytechnic Institute and State University traveled to nine sites, from northern Georgia to New Jersey, where beetles were well established. The researchers placed mesh bags around some adelgid-infested hemlock branches to keep predators away. They then monitored the bagged branches, as well as unprotected branches, from 2014 to 2018.

One indicator, the number of adelgid egg sacs showing signs of beetle attack, was "about as good a start as we could have hoped for," Salom says. Overall, beetles appear to have snacked on 30% to 40% of egg sacs on unbagged branches, the researchers reported last month in *Biological Control*.

Still, Salom and other researchers are hardly doing a victory lap. The beetle is not yet eating enough adelgids to reduce total populations, they found. The pest produces two generations per year, one in winter and one in spring. But the beetles only go after adelgids in winter. That means the adelgids can rebound to high numbers in spring and summer, when the beetles burrow into the soil and go dormant.

In addition, the beetle has not become well established in the northern forests that are now on the front lines of the adelgid invasion, says entomologist Mark Whitmore of Cornell University. Whitmore and others are working to establish two species of insects from western North America, called silverflies, which are known to eat adelgids in spring and summer. If successfully established in eastern forests, silverflies could become "one of the most important [adelgid] predators in the north," Whitmore says. In combination with beetles, the flies could also help control the adelgid's spring generation farther south, Salom adds.

Despite the headwinds facing the beetle project, Ellison says success could help set a precedent. Most efforts to control invasive insects have relied on parasites and parasitoids, which lay eggs and complete their life cycles in or on the target species. If Salom's team can build on its result to show that the beetles can really save trees, it would mark a rare example of successful biocontrol using a predatory insect.

Before biocontrol agents are deployed on a large scale, researchers typically conduct studies evaluating the potential risks and benefits (*Science*, 10 August 2018, p. 542). And Salom says his team rigorously vetted the beetle in the lab before releasing it. Still, releasing alien species in an ecosystem can have unforeseen consequences, Ellison notes. Introduced *L. nigrinus* beetles, for example, are now mating with related beetles native to eastern forests; about 2% of the beetles Salom's team recovered were such hybrids. That low rate suggests hybridization is likely "of little long-term concern," says Lynne Rieske-Kinney, an entomologist at the University of Kentucky.

Researchers are planning a follow-up study to determine whether the beetles are actually improving hemlock health, Salom says. For now, however, he's savoring a morsel of success. "I've been working on this for 22 years," he says. "It really feels good to have some positive data." ■

Gabriel Popkin is a writer in Mount Rainier, Maryland.

PHOTO: ROBBIE FLOWERS/VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

Published by AAAS

PROFESSION

‘Diversity statements’ divide mathematicians

Hundreds choose sides both for and against requiring such statements in academic hiring

By **Michael Price**

When Abigail Thompson, chair of the math department at the University of California (UC), Davis, penned an essay in the December 2019 issue of the *Notices of the American Mathematical Society* criticizing mandatory “diversity statements” in university hiring, simmering frictions in math boiled over. Researchers rushed to author op-eds and joint public letters both supporting and opposing Thompson. The reactions reflect a tension between mathematicians who see efforts to promote diversity as an intrusion of politics into research, and those who see opening their field to historically marginalized communities as the surest way to advance research. As befits the field, each side claims numerical data support their view.

Academic math skews overwhelmingly white and male. According to American Mathematical Society data, women made up 30% of new tenure-track hires in 2018. Data for ethnicity are harder to come by, but black and Hispanic faculty are rare. According to a 2017 National Science Foundation diversity report, black people are earning fewer degrees in math today than 20 years ago.

To improve diversity in all fields, eight of the 10 UC campuses as well as an increasing number of other universities across the country require faculty job applicants to submit a statement explaining steps they have taken to enhance diversity and equity, as well as how they will support students and colleagues from diverse backgrounds as a university employee. At UC, hiring committees consider these statements alongside teaching experience, research, and university and public service.

“This is a contentious issue whose discussion has been suppressed,” Thompson tells *Science*. Her essay applauds inclusivity efforts of recent decades, including encouraging people from diverse backgrounds to pursue a math career. But she calls diversity statements a “political litmus test” that she compares to the McCarthyist “loyalty oaths” faculty members were asked to sign in the 1950s disavowing association with communist thought. “Whatever our views on diversity and how it can be achieved, mandatory diversity statements are equally misguided,” she wrote.

Thompson says she has received “strong support” from colleagues at UC Davis and elsewhere. But to Chad Topaz, a mathematician at Williams College, requiring such diversity statements isn’t political; it’s a recognition that math generates better research when everyone can participate. “We want math to be the best it can be, so we have to make it accessible to everybody.”

Carrie Diaz Eaton, a computational scientist at Bates College, agrees. Building diversity, she says, requires “explicitly encouraging our junior colleagues and mentoring and cultivating their talents and enthusiasm for this work.”

ILLUSTRATION: DAVIDE BONAZZI/SALZMAN ART

Bands of mathematicians, including Topaz and Eaton, wrote and cosigned public letters responding to Thompson’s essay, 17 of which were published in December 2019 on the *Notices* website. Three of the letters each gained hundreds of signatures.

One heavily criticized Thompson’s essay and spoke in favor of diversity statements and other attempts to boost diversity. A second letter supported Thompson, calling her courageous and agreeing that making diversity statements mandatory is a mistake. A third struck a middle ground, supporting Thompson’s right to voice her opinion and calling for more research into the effectiveness of diversity statements.

Topaz, Eaton, and several other like-minded mathematicians then organized a quick research project, hiring contractors to look up publicly available information about the signatories. In a preprint on SocArXiv, they revealed that cosigners of the two pro-Thompson letters were overwhelmingly male (79% and 85%, respectively) and white (89% for both). The majority (88% and 60%) had tenure. By contrast, cosigners of the letter in favor of diversity statements were evenly split between men and women, and about 80% were white. And only one-quarter were tenured, with another 45% or so on the tenure track. In other words, Topaz says, the mathematicians who support mandating diversity statements “look a lot like the demographics of America” and represent math’s future.

Carina Pamela Curto, a mathematician at Pennsylvania State University, University Park, sees a different story in the numbers. She describes herself as pro-diversity, but she signed the letter calling required diversity statements a mistake. When hiring committees prioritize diversity, candidates’ other skills are necessarily devalued, she says. Especially for women and candidates from underrepresented races and ethnicities, she adds, spending time and effort developing diversity credentials can distract from doing research.

She and graduate student Joshua Paik reanalyzed Topaz’s data and found that 71% of the three letters’ female signatories who are tenured professors at research-intensive institutions signed one or both of the letters supporting Thompson. To her, the finding indicates that once women serve on hiring committees, their enthusiasm for diversity statements often cools.

However, Curto describes Thompson’s comparisons to McCarthyism as “totally counterproductive,” saying the piece “polarized the community in unfortunate ways.”

The community will have a chance to work things out face-to-face this week in Denver at the annual Joint Mathematics Meetings, which features dozens of sessions dedicated to diversity, equity, and social justice in math. ■

Published by AAAS

FEATURES

TRANSPARENCY ON TRIAL

Many clinical trial results aren't posted publicly, as U.S. law requires—and a promised crackdown has fizzled

By **Charles Piller**

For 20 years, the U.S. government has urged companies, universities, and other institutions that conduct clinical trials to record their results in a federal database, so doctors and patients can see whether new treatments are safe and effective. Few trial sponsors have consistently done so, even after a 2007 law made posting mandatory for many trials registered in the database. In 2017, the National Institutes of Health (NIH) and the Food and Drug Administration (FDA) tried again, enacting a long-awaited "final rule" to clarify the law's expectations and penalties for failing to disclose trial results. The rule took full effect 2 years ago, on 18 January 2018, giving trial sponsors ample time to comply. But a *Science* investigation shows that many still ignore the requirement, while federal officials do little or nothing to enforce the law.

Science examined more than 4700 trials whose results should have been posted on the NIH website ClinicalTrials.gov under the 2017 rule. Reporting rates by most large pharmaceutical companies and some universities have improved sharply, but performance by many other trial sponsors—including, ironically, NIH itself—was lackluster. Those sponsors, typically either the institution conducting a trial or its funder, must deposit results and other data within 1 year of completing a trial. But of 184 sponsor organizations with at least five trials due as of 25 September 2019, 30 companies, universities, or medical centers never met a single deadline (see https://scim.ag/ctgov). As of that date, these habitual violators had failed to report any results for 67% of their trials and averaged 268 days late for those and all trials that missed their deadlines. They included such eminent

ILLUSTRATION: DAVIDE BONAZZI/SALZMAN ART

Published by AAAS

institutions as the Harvard University–affiliated Boston Children's Hospital, the University of Minnesota, and Baylor College of Medicine—all among the top 50 recipients of NIH grants in 2019.

The violations cover trials in virtually all fields of medicine, and the missing or late results offer potentially vital information for the most desperate patients. For example, in one long-overdue trial, researchers compared the efficacy of different chemotherapy regimens in 200 patients with advanced lymphoma; another—nearly 2 years late—tests immunotherapy against conventional chemotherapy in about 600 people with late-stage lung cancer.

Other leading NIH grantees did only slightly better in *Science*'s analysis based on data collected from the TrialsTracker website of the University of Oxford, which automatically mines ClinicalTrials.gov information. The University of Texas MD Anderson Cancer Center and the Mayo Clinic both failed to report results on time, or at all, in about two-thirds of their trials. Yale University failed to do so in 84% of its trials. NIH's own institutes also had a bad record. They are directly responsible for reporting results when they sponsor studies done by agency staff or some grantees, and the top four NIH institute sponsors, taken together, reported results late or not at all in more than six of every 10 trials *Science* looked at.

Contacted for comment, none of the institutions disputed the findings of this investigation. In all 4768 trials *Science* checked, sponsors violated the reporting law more than 55% of the time. And in hundreds of cases where the sponsors got credit for reporting trial results, they have yet to be publicly posted because of quality lapses flagged by ClinicalTrials.gov staff (see sidebar, p. 243).

Although the 2017 rule, and officials' statements at the time, promised aggressive enforcement and stiff penalties, neither NIH nor FDA has cracked down. FDA now says it won't brandish its big stick—penalties of up to $12,103 a day for failing to report a trial's results—until after the agency issues further "guidance" on how it will exercise that power. It has not set a date. NIH said at a 2016 briefing on the final rule that it would cut off grants to those who ignore the trial reporting requirements, as authorized in the 2007 law, but so far has not done so.

Many scientists who conduct clinical trials, and their sponsors or funders, have downplayed concerns about late or missing results in ClinicalTrials.gov. Researchers, doctors, and patients can instead learn about trial outcomes from peer-reviewed publications, they say. But thousands of trials are never published, particularly when they find treatments ineffective, history has shown.

ClinicalTrials.gov also uses a common format, allowing comparisons of results across trials that journal articles rarely make possible. Doctors, researchers, and potential trial participants rely on the site, to judge from its 215 million monthly page views.

Deborah Zarin, a physician at Brigham and Women's Hospital and Harvard who headed ClinicalTrials.gov between 2005 and 2018, says the *Science* findings show failures of the research culture, FDA, and NIH. "If this was a priority for the leadership of NIH, then they could ensure that high-quality, timely reporting happened all of the time," says Zarin, an NIH-paid research consultant for the database. "You can set up processes so trial reporting is an expectation. You can't pass 'go' and collect $200 until this is done."

Zarin, who works in a program to advance clinical research, adds that the problem persists because "reporting to ClinicalTrials.gov is frequently seen by sponsors, funders, and trialists as an annoying administrative and perhaps legal burden, not a scientific imperative. Human nature being what it is, people follow the requirements when forced to do so."

NIH and FDA officials do not seem inclined to apply that pressure. Lyric Jorgenson, NIH deputy director for science policy, says her agency has been "trying to change the culture of how clinical trial results are reported and disseminated; not so much on the 'aha, we caught you,' as much as getting people to understand the value, and making it as easy as possible to share and disseminate results." To that end, she says, ClinicalTrials.gov staff have educated researchers about the website and improved its usability.

As for FDA, Patrick McNeilly, an official at the agency who handles trial enforcement matters, recently told an industry conference session on ClinicalTrials.gov that "FDA has limited resources, and we encourage voluntary compliance." He said the agency also reviews reporting of information on ClinicalTrials.gov as part of inspections of trial sites, or when it receives complaints.

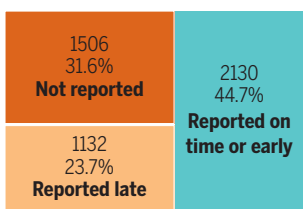
McNeilly declined an interview request, but at the conference he discounted violations of ClinicalTrials.gov reporting requirements found by journalists and watchdog groups. "We're not going to blanketly accept an entire list of trials that people say are noncompliant," he said. Such determinations require "nonpublic information" submitted to the agency by trial sponsors. In response to *Science*'s findings, a spokesperson said an absence of posted results on ClinicalTrials.gov did not mean a trial sponsor has broken the 2007 law.

Yet that law and the 2017 final rule detail only a few exemptions that would allow trial sponsors to withhold results on the basis of nonpublic information. The very few registered trials that qualify for those exemptions are not flagged as violators by TrialsTracker or in *Science*'s analysis.

CONGRESS APPROVED the creation of ClinicalTrials.gov in 1997, after allegations that patients were harmed because companies withheld evidence showing their medicines were ineffective or hazardous. A widely cited case involved the GlaxoSmithKline antidepressant Paxil (paroxetine). According to legal filings and a report in *The BMJ*, the firm held secret data showing that in clinical trials

Missed deadlines

Among more than 4700 clinical trials examined by *Science*, less than 45% had their results reported early or on time to ClinicalTrials.gov.

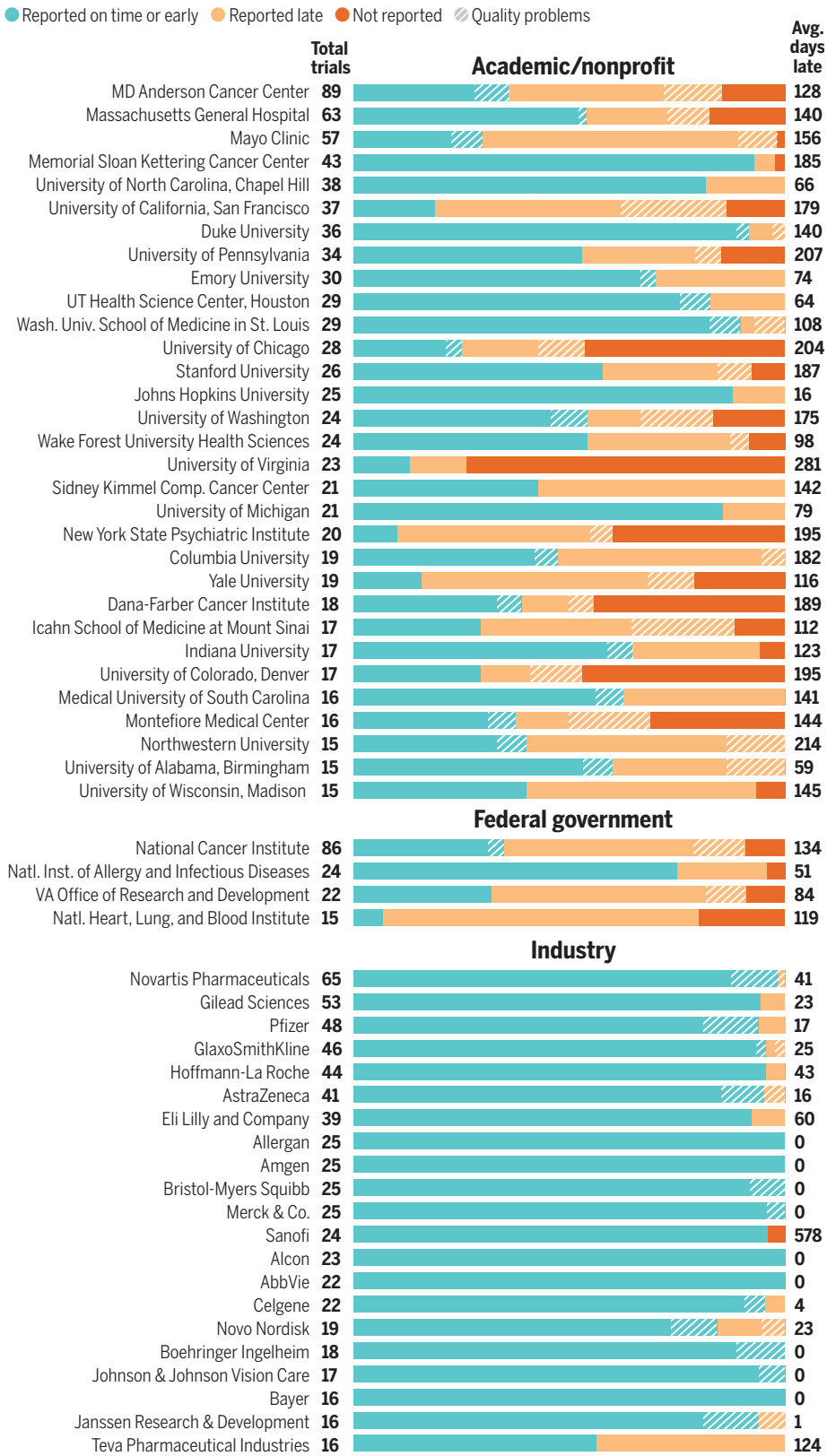


CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) CLINICALTRIALS.GOV/TRIALSTRACKER

Published by AAAS

Reporting problems

Science analyzed ClinicalTrials.gov records of all clinical trials with results required to be reported between 18 January 2018 and 25 September 2019. The chart, covering trial sponsors with 15 or more results due in that window or reported early, shows that some results deposited were not posted due to quality lapses. It also highlights that pharma's record has been markedly better than that of academia and the federal government.



CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) CLINICALTRIALS.GOV/TRIALSTRACKER

the drug was ineffective and caused suicidal thoughts in teenagers, yet encouraged doctors to prescribe it for young people.

Registration was only required initially for trials of treatments for serious or life-threatening diseases. But the 2007 law, the Food and Drug Administration Amendments Act, required sponsors to register a much broader range of trials within 21 days of enrolling the first patient, and to post summary results, adverse events, and other data to ClinicalTrials.gov within 1 year of collecting the last patient data. Although many trials, such as industry-sponsored early-stage evaluations of drug safety, are exempt from reporting, about 326,000 have been registered, and results have been posted for more than 40,000.

Yet until 2015, even the most active investigators at clinical research institutions treated the law more as a suggestion—not surprising given that the government enforced no penalties and did not publicly identify violators. A report on the news website *STAT* by this author and Talia Bronshtein first drew significant attention to specific trial sponsors—companies, government agencies, universities, and individuals—that routinely ignored reporting requirements. It sparked immediate improvement, according to NIH. (Those same authors documented some of that improvement in a 2018 *STAT* article.)

At a 2016 press briefing, NIH and FDA rolled out the final rule, aimed at boosting even greater compliance with the 2007 law. It took effect in January 2017, with first deadlines for results, and ostensibly enforcement, 1 year later. Then–FDA Commissioner Robert Califf said it would thereafter "be pretty hard to hide that you are doing a clinical trial or hide the result." FDA, he vowed, was finally prepared, if necessary, to enforce the daily $10,000 penalty for noncompliance allowed under the law. (Adjusted for inflation, that figure recently rose above $12,000.)

"I don't think anybody wants to be on the wall of shame," NIH Director Francis Collins said at the press event, promising that NIH would publicly flag reporting violations on ClinicalTrials.gov itself.

"We are serious about this," Collins said, threatening for the first time to enforce provisions of the 2007 law that allow NIH to rescind funding to grantees who violate the statute. "It's hard to herd cats, but you can … take their food away," he said. "This is about maintaining the trust that we have with participants in clinical trials. … If we fail to live up to that expectation, then that is an ethical failure."

Three years later, TrialsTracker conservatively estimates that FDA could have collected more than $6 billion in ClinicalTrials.gov penalties so far. The agency has yet to demand a single dollar. And despite

Published by AAAS

Gaming the system

ClinicalTrials.gov relies on the organizations that fund or run clinical trials to register studies and report accurate results on time. Many miss the legal reporting deadlines (see main story, p. 240). But even those that comply often do so with substandard effort. They face no penalty: Sponsors are credited with fulfilling their legal obligations on the date they file the faulty data.

Deborah Zarin, who headed the National Institutes of Health data site from 2005 to 2018, recently reported in *JAMA Internal Medicine* that typical lapses in reported results included showing data on more participants than were enrolled and providing inconsistent units of measurement—rendering the results incoherent. ClinicalTrials.gov analysts reject such results until they pass basic quality-control standards—a process than can take many months. *Science* has found that one of every seven trials credited with reporting results shows no posted data as a result of quality-control reviews. Results for many other trials were posted only after multiple rejections.

Similar quality issues often delay an earlier step: registering a trial at the outset. According to Zarin, ClinicalTrials.gov has historically rejected half of all initial registrations because of a basic flaw: unclear primary outcome measures—what a trial seeks to examine. Often, rejections stem from "sloppiness or lack of rigor," she says. But she has heard of companies that want to appear to comply with ClinicalTrials.gov's legal mandate but deliberately file shoddy trial registrations—likely to be rejected and therefore delay public posting—to temporarily protect information they regard as proprietary.

Jennifer Miller, an ethicist at Yale University who created the Good Pharma Scorecard to rank drug companies on ethics and transparency, says some trial sponsors go even further. In a recent study in *The BMJ*, Miller and colleagues examined 19 drugs approved by the U.S. Food and Drug Administration (FDA) in 2015, all from large companies. In only 10 cases were all the trials provided to FDA for its decisions registered on ClinicalTrials.gov. When studies aren't registered, a failure to report results is impossible to detect. "It's easy to game the system," Miller says. "Just don't register your trial."

"Minimal compliance with the law should be a given," she adds. But trial sponsors should do more than just go through the motions of registering their studies and posting results, Miller says. Evidence-based medicine demands complete and careful reporting: "Ethics is not minimal compliance with the law." —*C. P.*

more than 2600 trials for which results are overdue or were filed late, NIH has yet to withhold a single grant as a result or post a single violation notice on ClinicalTrials.gov. No "wall of shame" exists.

"Public-facing websites run by the government should be accurate. That's not asking much," Senator Chuck Grassley (R–IA), who advocated for the 2007 law, wrote in an email after reviewing a summary of the *Science* findings. "It's a question of basic management and agency competence. The government has a duty to police its work product, especially because the public trusts .gov websites will be accurate and reliable."

To physician Ben Goldacre, who directs the Oxford program behind TrialsTracker, "The lack of urgency is really troubling."

IN A RECENT ARTICLE in *The BMJ*, Goldacre and colleagues highlighted a long-running MD Anderson trial as an example of what's at stake when clinical research results go AWOL. Started in 1999, the trial tested a specialized hormone therapy in patients who had surgery for prostate cancer and faced a high risk of recurrence. The treatment might reduce that risk significantly, but other trials suggested it had serious side effects, including reduced bone density, sexual dysfunction, and greater risk of diabetes and cardiovascular events. When *Science* collected the ClinicalTrials.gov data, the MD Anderson trial's results were overdue by 1 year and 8 months without explanation, and no journal appears to have published them. Doctors and prostate cancer patients weighing the most appropriate treatment have been left in the dark.

Stephen Hahn, who served as chief medical officer at MD Anderson until last month, when he became the new FDA commissioner, was unavailable for comment. An MD Anderson spokesperson said the center "believes in transparency," and is making "every effort to comply" with trial reporting rules, but did not respond to a question on the missing cancer trial data.

Mayo, Yale, the University of Minnesota, Baylor, and Boston Children's, which have similarly poor reporting records, all said via email that they, too, were committed to fulfilling ClinicalTrials.gov requirements. Mayo said it deployed dedicated staff to assist researchers. Yale noted it had registered hundreds of trials exempt from reporting requirements and added a layer of review to help ensure compliance. The University of Minnesota recently created a monitoring system and has made rapid progress on reporting, according to a spokesperson. Baylor said it planned to centralize trial monitoring, as "a top priority." Boston Children's said it was "committed to achieving 100% compliance." It submitted data for one long-overdue study after being contacted by *Science*, but four others remain in apparent violation as of the end of 2019.

The slow, apparently forgiving regulatory approach favored by NIH and FDA will never force such organizations into full compliance, some advocates of clinical trial transparency say. "In an era when every restaurant is obliged to publish on its front door the hygiene rating in their kitchen, we're seriously not saying whether a trial that costs millions of dollars has broken the law and its obligation to … patient participants by failing to report its results?" Goldacre asks. "That seems like extraordinary special treatment for clinical trialists."

Select institutions have made serious efforts to comply. Twenty big pharma companies met all reporting requirements under the 2017 rule and some major academic centers improved sharply compared with data collected in 2017 (as detailed in the second *STAT* investigation). Memorial Sloan Kettering Cancer Center, Duke University, and Johns Hopkins University—poor performers in 2017—complied with the law in nearly all their registered trials covered by the new rule. Johns Hopkins added staff to track and assist on reporting and to identify "problem records." And it enlisted university executives to crack down on recalcitrant investigators, according to Anthony Keyes, a clinical research manager there.

But such good performance shouldn't be an exception, Harvard's Zarin says. "Further public accountability of the trialists, but also our government organizations, has to happen. One possibility is that FDA and NIH will be shamed into enforcing the law. Another possibility is that sponsors will be shamed into doing a better job. A third possibility is that ClinicalTrials.gov will never fully achieve its vital aspirations." ■

This story was supported by the *Science* Fund for Investigative Reporting.

Published by AAAS

PALEONTOLOGY

Eat and listen—how chewing and hearing evolved

Mammalian middle ear bones separated from the jaw of vertebrate ancestors

By **Julia A. Schultz**

In rare cases, fossils function like snapshots, capturing key moments of evolution. One of these rare cases is *Origolestes lii*, a stem therian mammal from the Yixian Formation (China) described by Mao *et al.* (*1*) on page 305 of this issue. The study provides a glimpse of a fundamental moment in mammalian evolution—the separation of middle ear elements from the Meckelian cartilage of the lower jaw, which provided freedom to the system to improve both chewing function and hearing ability.

Chewing is specific to mammals; no other vertebrates use elaborate repeated jaw movements to process food before swallowing. Theria, the group all living placentals and marsupials descend from (see the figure), initially utilized upward but slightly transverse (sideward) movements of the lower jaw for processing food. Within this group, versatile chewing movements evolved for efficient

Rheinische Friedrich-Wilhelms-Universität Bonn, Institut für Geowissenschaften, Bonn, Germany.
Email: jaschultz@uni-bonn.de

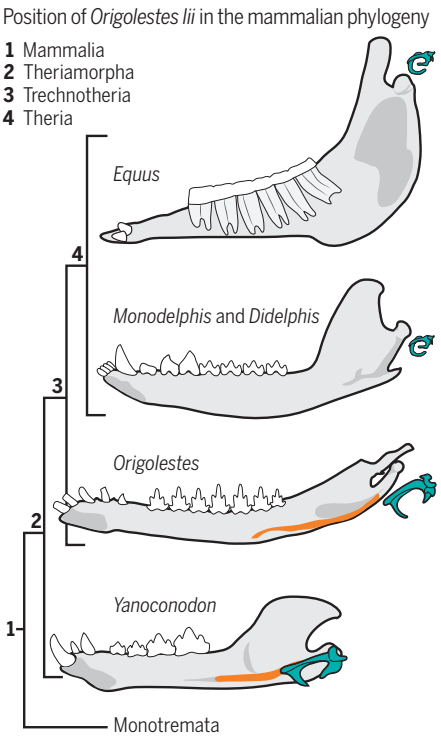
ILLUSTRATION: CHUANG ZHAO

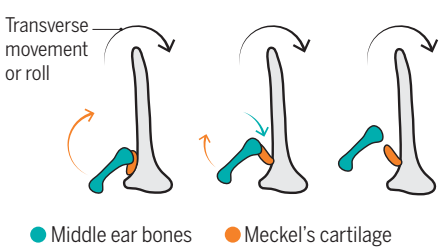
Published by AAAS

Specimens of *Origolestes lii* were discovered in the Lujiatun beds of the Yixian Formation, Liaoning Province, China. The illustration shows that the animals died at rest, a condition similar to those found for other vertebrates from the same locality, including dinosaurs.

Evolution of the therian dentary and middle ear

The connection of the middle ear elements (teal) to the lower jaw by means of the ossified Meckelian cartilage (orange) is lost.





breakdown of all kinds of food by intensifying the transverse movement and integrating long-axis rotation (that is, rolling) (*2*). Chewing reduces the size of food to facilitate access to nutrients. The smaller the particles, the larger the surface available to gut microbiota, and acquiring the ability to digest a new diet is a fundamental driver for the evolution of new species (*3*). Mammalian teeth play a key role in food processing because they function as tools for breaking food into pieces. In therian mammals, a typical form of molars evolved (the tribosphenic, or multifunctional molar) with a characteristic cusp and basin pattern. This molar type is considered a revolution in mammalian evolution. Its functionality widened the range of food resources (*4*), and that this tooth type is retained in the majority of omnivorous and insectivorous mammals living today speaks for its universal utility. It is consensus that the tribosphenic molar gave rise to the variety of tooth forms of extant mammals (*5*). However, the morphology combining cutting and crushing functions is only fully functional in combination with a transverse movement of the lower jaw.

By definition, mammals possess a single bone (dentary) as the tooth-bearing lower jaw, and multiple tiny ossicles in the middle ear (malleus, incus, and stapes). In mammalian ancestors (nonmammalian cynodonts), the lower jaw was composed of multiple bones (angular, quadrate, and articular) that transmitted sound by bone conduction through the so-called mandibular ear, in addition to its function in catching prey and chewing. In these early forms, the quadrate and articular formed the primary jaw joint, an articulation found in all jawed (gnathostome) vertebrates (*6*). Transitional taxa like *Morganucodon* had a double craniomandibular joint; this consisted of a primary jaw joint that coexisted with a derived secondary articulation between the condyle (rounded hinge area of a bone) of the lower jaw and the squamosal (bone part of the vertebrate skull of the cranium). In these early mammalian forms, the primary jaw joint was still involved in both feeding and transmitting sound to the inner ear. The middle ear bones, still attached to the lower jaw and sitting in grooves on the inner side of the dentary (postdentary trough) were limited in their mobility. In such close connection, hearing and feeding functions interfered. Transverse shift or even rolling of the dentary bone was limited, because these movements would entail displacement of the middle ear elements inside the postdentary trough (see the figure). During the evolution of the mammalian feeding function, the secondary jaw articulation prevailed. The middle ear bones shrank and shifted posteriorly, and the innovative secondary jaw joint allowed more powerful jaw movements and effective control of the dentary during chewing (*7*).

Homology of the primary jaw joint bones and the mammalian middle ear elements is a classic case study of embryology, and similarities were identified in the 19th century. Supporting tissues, like the Meckelian and palatoquadrate cartilages of the first vertebrate pharyngeal arch, are embryonic precursors to the articular and quadrate of nonmammalian gnathostome vertebrates and malleus and incus of living mammals (*8*). Genetic mechanisms controlling the formation of the mammalian jaw and middle ear became an area of interest in developmental biology in the past 25 years. A series of cellular and molecular drivers in the developmental program of vertebrate pharyngeal arches were identified, which also control formation and detachment of the mammalian middle ear ossicles from the lower jaw (*9*, *10*).

The complex developmental process of the detachment of the middle ear elements is well understood from an ontogenetic perspective, but timewise, the separation in the stem lineage of therians was not known from the fossil record. *O. lii* is the first fossil that pinpoints the moment of detachment of the middle ear ossicles from the Meckelian cartilage to Lower Cretaceous (~123 million years ago). Other stem therian representatives like *Yanoconodon allini* or *Maotherium asiaticus*, slightly older than *O. lii*, show separate

GRAPHIC: ADAPTED FROM (*1*) AND (*15*) BY N. CARY/*SCIENCE*

Published by AAAS

middle ear ossicles from the jaw but maintain a slender bony connection to the ossified Meckelian cartilage (*11*, *12*). In *O. lii*, the connection to the ossified Meckelian cartilage is lost and the middle ear ossicles lie separate, likely suspended from the cranium by soft tissue connections.

As stated by Mao *et al.*, *O. lii* is evidence for modular evolution of chewing and hearing. Decoupling the two functions was crucial to mammalian evolution to increase chewing efficiency and hearing ability. Decoupling and reducing the size of the middle ear ossicles proved advantageous for better sound transmission. The incorporation of multiple delicate and mobile elements into the middle ear cavity increased the hearing sensitivity of mammals, most notably to high-frequency sounds (*13*), which is considered an advantage for detecting insect prey. High-frequency singing in insects was already established by the mid-Jurassic, as known from cricket fossils that show preserved stridulatory files with cuticular teeth on their wings (*14*).

The complex process of detachment of the middle ear ossicles happened at least three times in the evolution of different mammalian lineages (Multituberculata, Monotremata, and Theria) (*15*). In all three groups, the middle ear ossicles became smaller and shifted to the basicranium. Only in therians did this process enhance transverse movement in the chewing process. Chewing movements became very diverse in this group and increased the chewing efficiency, a phenomenon realized to an extreme in modern herbivores such as horses and camels. In connection with the modification of the dentition (for example, high-crowned teeth) complex feeding adaptations were possible within this group only by decoupling hearing and feeding functions. ■

REFERENCES AND NOTES

1. F. Mao *et al.*, *Science* **376**, 305 (2020).
2. D. M. Grossnickle, *Sci. Rep.* **7**, 45094 (2017).
3. R. E. Ley *et al.*, *Science* **320**, 1647 (2008).
4. A. V. Lopatin, A. O. Averianov, *Science* **313**, 1092 (2006).
5. A. R. Evans, G. D. Sanson, *Biol. J. Linn. Soc. London* **78**, 173 (2003).
6. E. F. Allin, J. A. Hopson, in *The Evolutionary Biology of Hearing*, D. B. Webster, A. N. Popper, R. R. Fay, Eds. (Springer, 1992), pp. 587–614.
7. A. W. Crompton, W. L. Hylander, in *The Ecology and Biology of Mammal-like Reptiles*, N. Hotton III, P. D. MacLean, J. J. Roth, E. C. Roth, Eds. (Smithsonian Institution Press, 1986), pp. 263–282.
8. M. Takechi, S. Kuratani, *J. Exp. Zool.* **314B**, 417 (2010).
9. T. Kitazawa *et al.*, *Nat. Commun.* **6**, 6853 (2015).
10. D. J. Urban *et al.*, *Proc. R. Soc. London Ser. B* **284**, 20162416 (2017).
11. Z.-X. Luo *et al.*, *Nature* **446**, 288 (2007).
12. Q. Ji *et al.*, *Science* **326**, 278 (2009).
13. S. Puria, C. Steele, *Hear. Res.* **263**, 183 (2010).
14. J. J. Gu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 3868 (2012).
15. Z.-X. Luo, *Annu. Rev. Ecol. Evol. Syst.* **42**, 355 (2011).

10.1126/science.aba3808

CELL BIOLOGY

Water loss regulates cell and vesicle volume

Active control of vesicle water content ensures that cell size is maintained when extracellular fluid is taken up

By **Jason S. King** *and* **Elizabeth Smythe**

When cells take up extracellular fluid by endocytosis, they internalize a considerable proportion of the cell volume quickly and yet maintain their volume and ionic composition. This is particularly striking in the case of macropinocytosis, which is the bulk uptake of extracellular fluid. Through this pathway, macrophages can be stimulated to internalize ~25% of their cellular volume per hour into large vacuoles known as macropinosomes. An intriguing question is how cells and organelles are able to maintain their size while internalizing such large volumes. On page 301 of this issue, Freeman *et al.* (*1*) reveal a molecular mechanism underpinning homeostatic regulation of cell size. They demonstrate that newly formed macropinosomes rapidly lose volume by osmosis driven by two-pore channel (TPC)–mediated outflow of sodium ions. This reduces hydrostatic pressure within the macropinosome, facilitating the extension of tubules from the macropinosome surface and recycling of membrane lipids and proteins back to the cell surface.

Cells need to continuously take up nutrients as well as signaling molecules from their surroundings. Clathrin-dependent and -independent endocytic pathways minimize uptake of extracellular fluid by generating small vesicles with a high surface area/volume ratio, whereas macropinocytosis allows cells to engulf large volumes of extracellular fluid. This is especially important in immune cells, because it enables surveillance of the extracellular environment (*2*). Moreover, cancer cells and many free-living protists use this pathway to take up essential nutrients, which allows them to survive and proliferate (*3*, *4*).

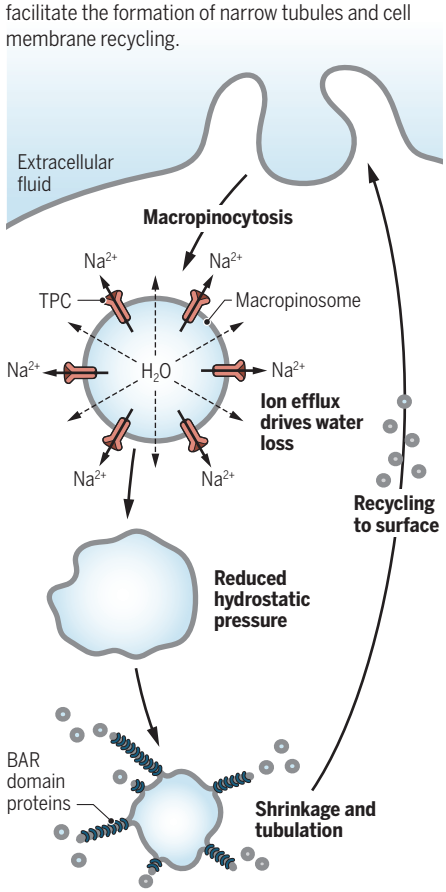
A universal feature of macropinosomes is that they shrink as they mature, allowing cells to resolve the large volumes internalized. Freeman *et al.* show that specific loss of sodium and chloride ions from macropinosomes is key for this shrinkage. They found that this is carried out by TPCs and is required for immune function in mice, demonstrating its importance in vivo.

TPCs are transmembrane channels that are specific for sodium transport and are activated by phosphatidylinositol 3,5-bisphosphate [$PI(3,5)P_2$] (*5*), a low-abundance lipid found on endosomes and lysosomes, the digestive organelles of the cell. $PI(3,5)P_2$ is generated by the enzyme phosphatidylinosi-

Centre for Membrane Interactions and Dynamics, Department of Biomedical Science, University of Sheffield, Sheffield S10 2TN, UK. Email: e.smythe@sheffield.ac.uk

Shrinking vesicles

Macropinosome formation allows cells to take in large volumes of extracellular fluid. To maintain cell size, macropinosomes need to lose fluid. Two-pore channels (TPCs) mediate efflux of sodium ions, which drives water loss through osmosis. This allows the recruitment of BAR domain proteins, which facilitate the formation of narrow tubules and cell membrane recycling.



GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

Published by AAAS

tol 3-phosphate 5-kinase (PIKFYVE), which is essential for correct processing of macropinosomes, as well as processing of vacuoles involved in digesting bacteria engulfed by phagocytosis (phagosomes) and recycled intracellular material (autophagosomes) (*6*–*9*). TPCs are not the only channels regulated by PI(3,5)P_2, and other researchers have previously shown that PIKFYVE is required for phagosome maturation in macrophages, at least partially through activation of the lysosomal cation channel mucolipin 1 (TRPML1) (*6*, *10*). Whether these two mechanisms are complementary or represent cell type–specific differences is unclear, but the regulation of vesicle shrinkage via lipid-gated ion channels appears to be a general mechanism. The use of TPCs to drive macropinosome shrinkage in mammalian macrophages is likely attributable to the extremely high sodium concentrations found in their extracellular fluid. Other channels may perform a similar role in eukaryotes living in different ionic environments.

A consequence of the rapid TPC-driven reduction in lumenal sodium is that water is also lost from the macropinosome by osmosis. Freeman *et al*. propose a mechanism by which the resultant drop in lumenal hydrostatic pressure promotes the tubulation and recycling of the macropinosome membrane. They show that BAR domain proteins, which can sense, stabilize, and sometimes induce membrane tubulation (*11*), are recruited in an osmotically sensitive manner. Therefore, in the presence of BAR domain–containing Myc box–dependent interacting protein 1 (BIN1), osmotic shrinkage facilitates the formation of narrow tubules (see the figure). It remains unclear whether the BAR domain proteins bind because they recognize curvature from small shrinkage-induced wrinkles, or because they respond to lowering of membrane tension.

Tubulation provides a mechanism to selectively remove and recycle surface proteins from vesicles while retaining lumenal contents. This is a common mechanism of endocytic recycling of membrane constituents (*11*). It is particularly important for macropinocytosis, which nonspecifically internalizes cell surface proteins such as receptors. Maintaining surface expression of such proteins is essential for normal cell functions. Surface proteins are therefore rapidly recycled from macropinosomes after internalization, preventing their degradation with the rest of the internalized material (*12*). Freeman *et al*. show that this recycling is sensitive to TPC-mediated ion flux.

How general is this mechanism for other vesicles? Although changes in volume in other endocytic pathways are smaller, internalized solvent still needs to be expelled. Any time a vesicle forms a tubule, proportionally more surface area is used than lumenal volume. To maintain hydrostatic pressure, tubulation must be accompanied by loss of lumenal content. Given the relatively small volumes that are internalized by other endocytic pathways and the continued delivery and retrieval of fluid from early endosomes, considerable shrinkage is unlikely to be observable. Nonetheless, BAR domain proteins mediate tubulation and recycling from nearly all endocytic pathways, implying that osmosis-mediated shrinkage may be of general importance. A more general role for osmotic regulation after clathrin-mediated endocytosis is also suggested by defects in recycling of integrins and transferrin receptors in nonmacropinocytic fibroblasts (*1*). Although the activity of TPCs is up-regulated in tissues with high macropinocytic activity, they are also expressed in cells such as fibroblasts where they are found on lysosomes, which also have high sodium concentrations (*13*). It has therefore been proposed that TPCs regulate lysosomal sodium concentrations to maintain acidification.

It is striking that both TPCs and TRPML1 are regulated by mechanistic target of rapamycin (mTOR) (*10*, *13*), a nutrient-sensitive signaling protein that responds to environmental cues to control cell growth (*14*). When nutrients are in abundance, mTOR phosphorylates TPCs, keeping them closed, whereas under starvation conditions, channels are constitutively open. mTOR signaling is frequently dysregulated in cancer and diabetes, and the cross-talk between mTOR and TPCs has suggested intriguing therapeutic possibilities (*15*). Given the key role of macropinocytosis in feeding cancer cells, it will be fascinating to explore how dysregulated mTOR signaling in cancer cells affects TPC-mediated macropinosomal shrinkage. ■

REFERENCES AND NOTES

1. S. A. Freeman *et al*., *Science* **367**, 301 (2020).
2. M. A. West *et al*., *Science* **305**, 1153 (2004).
3. C. Commisso *et al*., *Nature* **497**, 633 (2013).
4. J. S. King, R. R. Kay, *Philos. Trans. R. Soc. B* **374**, 20180158 (2019).
5. X. Wang *et al*., *Cell* **151**, 372 (2012).
6. R. M. Dayam *et al*., *Traffic* **16**, 1010 (2015).
7. J. de Lartigue *et al*., *Traffic* **10**, 883 (2009).
8. J. Castonguay *et al*., *Sci. Rep.* **7**, 10038 (2017).
9. C. M. Buckley *et al*., *PLOS Pathog.* **15**, e1007551 (2019).
10. S. Krishna *et al*., *Dev. Cell* **38**, 536 (2016).
11. M. Simunovic *et al*., *Annu. Rev. Cell Dev. Biol.* **35**, 111 (2019).
12. C. M. Buckley *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **113**, E5906 (2016).
13. C. Cang *et al*., *Cell* **152**, 778 (2013).
14. K. J. Condon, D. M. Sabatini, *J. Cell Sci.* **132**, jcs222570 (2019).
15. C. Grimm *et al*., *Front. Pharmacol.* **8**, 45 (2017).

ACKNOWLEDGMENTS

J.S.K. is supported by a Royal Society Fellowship.

10.1126/science aba3623

IMMUNOLOGY

Enforcing T cell innocence

Inhibitory signaling on T cells maintains "naïvety" to self-antigen

By **Chrysothemis Brown**[1,2] *and* **Alexander Y. Rudensky**[1,2]

The adaptive immune system, with T and B lymphocytes being its two principal cell types, enables efficient protection from countless pathogens. T lymphocytes emerge from the thymus as quiescent (noncycling) naïve cells and display an immense repertoire of clonally distributed T cell receptors (TCRs) that detect antigens and are generated as the result of a largely random process of somatic gene rearrangement in thymic precursor cells. On page 264 of this issue, ElTanbouly *et al*. (*1*) show that naïve T cell quiescence is actively regulated upon their egress from the thymus and that this may

> **"...naïve T cell quiescence is actively maintained to prevent inappropriate activation and ensuing autoimmunity."**

be critical for constraining self-reactive T cells and preventing autoimmunity.

TCRs recognize antigens in the form of peptide fragments, bound to major histocompatibility complex (MHC) molecules. These peptides are generated intracellularly either from proteins of foreign origin, including pathogens, or from endogenous "self" proteins. During thymic differentiation, an intricate selection process eliminates naïve T cells with a strong affinity for self-peptide–MHC complexes. Upon egress from the thymus, these T cells, identified by low expression of the cell adhesion molecule CD44 and high expression of the lymph node homing receptor CD62L, continue to survey peptide-MHC complexes displayed on the surface of antigen-presenting cells.

[1]Howard Hughes Medical Institute and Immunology Program, Sloan Kettering Institute, New York, NY 10065, USA. [2]Ludwig Center, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
Email: brownc10@mskcc.org; rudenska@mskcc.org

Published by AAAS

This continuous TCR "tickling" is thought to support the maintenance of peripheral naïve T cells (*2*), which undergo activation, clonal expansion, and differentiation into effector T cells when they encounter their cognate antigen during inflammation.

More than 100 years ago, the pioneering physician and scientist Paul Ehrlich noted that the immune system of higher organisms can turn against the organism, leading to the development of autoimmunity. Decades later, it was proposed that there are two signals for T cell activation—one transmitted through the TCR and a second transmitted through a costimulatory receptor (*3*, *4*). Discovery of CD28 as the principal costimulatory receptor, with CD80 and CD86 as its ligands, provided support for this model. The subsequent discovery that expression of these costimulatory molecules was directly linked to sensing of microbial ligands by evolutionary conserved pattern recognition receptors (for example, Toll-like receptors) provided the basis for a two-signal model of naïve T cell activation that integrates innate and adaptive receptor cues (*5*). This suggests that self-reactivity is limited in a "passive" manner through lack of a permissive second signal (CD80 or CD86).

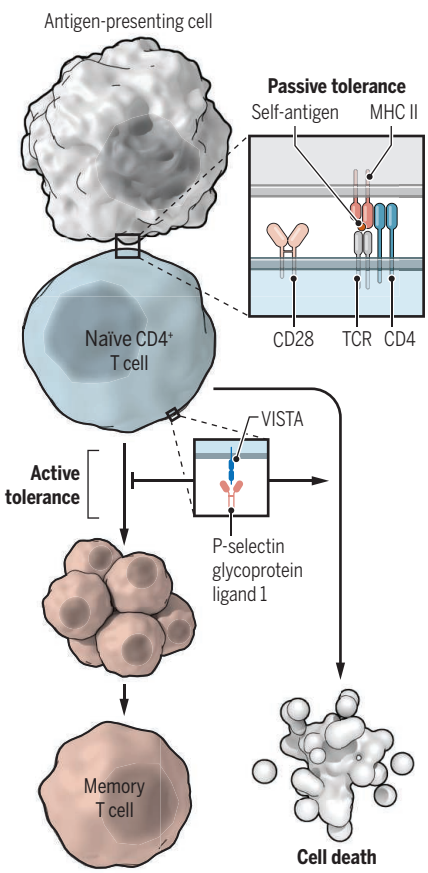
Another layer of regulation has been identified through the discovery of inhibitory receptors, CTLA-4 (cytotoxic T-lymphocyte–associated antigen 4) and PD-1 (programmed cell death protein 1), that are induced upon activation of T cells to dampen their responses (*6*). These mechanisms represent an "active" mode of reinforcement of self-tolerance through limiting TCR and CD28 signaling. The importance of this active pathway in limiting self-reactivity is demonstrated by genetic studies in mice as well as therapeutic targeting of these receptors in cancer, both of which can lead to the development of autoimmunity (*7*).

ElTanbouly *et al.* reveal a role for VISTA (V-type immunoglobulin domain-containing suppressor of T-cell activation), an inhibitory receptor that is homologous to the PD-1 ligand, PDL1 (*8*), in preventing activation of naïve T cells in the absence of inflammation. Mice deficient in VISTA have increased frequencies of effector T cells and, in some instances, develop autoimmunity (*9*, *10*). On myeloid cells, VISTA acts as an inhibitory ligand. However, T cells also have low expression of VISTA, where it may act as an inhibitory receptor (*11*). ElTanbouly *et al.* show that constitutive expression of VISTA by T cells in homeostasis acts as a brake to restrain inappropriate naïve T cell activation. This indicates that naïve T cell quiescence is actively maintained to prevent inappropriate activation and ensuing autoimmunity (see the figure).

Despite the checkpoints in place to prevent autoimmunity, a small proportion of naïve T cells spontaneously convert to memory-like CD4$^+$ T cells, possibly in response to self-antigen recognition (*12*). ElTanbouly *et al.* provide evidence that VISTA may serve as a critical T cell–intrinsic mechanism for preventing this "homeostatic" conversion of naïve to effector T cells. Using single-cell RNA sequencing, they identified a dominant cluster of naïve CD4$^+$ T cells in healthy mice distinguished by high expression of transcription factors known to regulate T cell quiescence (*13*). Although thymic naïve CD4$^+$ T cells also express VISTA, its loss in mice did not affect their phenotype until after they left the thymus. This suggests that VISTA acts to prevent T cell activation in response to signals received in the periphery.

Naïve T cell tolerance

TCRs recognize self-antigens bound to MHC II molecules. In passive tolerance, naïve T cells that recognize MHC II–self-antigen in the absence of costimulatory CD80 or CD86 become hyporeactive. In active tolerance, VISTA promotes naïve T cell death and prevents conversion to the memory phenotype after self-antigen recognition.



MHC II, major histocompatibility complex type II; TCR, T cell receptor; VISTA, V-type immunoglobulin domain-containing suppressor of T-cell activation.

Do these shifts in naïve T cell phenotypes require TCR signaling? Providing further support for a model whereby TCR engagement, potentially by self or innocuous non-self antigens, drives conversion of naïve T cells to memory-like CD4$^+$ cells in steady state, ElTanbouly *et al.* demonstrate that VISTA specifically suppresses clonal expansion of naïve T cells after engagement of the TCR in the absence of costimulation.

These findings point to active regulation of naïve T cell quiescence as a critical mechanism for the prevention of autoimmunity. The interacting partner for VISTA was recently identified as P-selectin glycoprotein ligand 1 (PSGL-1) (*14*), which is not only expressed on the surface of both myeloid and activated T cells but also on resting T cells, albeit at a lower amount. Further elaboration of this signaling pathway is needed to understand how VISTA regulates T cell quiescence. Understanding the signals that act upstream of VISTA will be important for harnessing this immunemodulatory pathway: either enhancing naïve T cell responses in settings of vaccination or suppressing T cell activation during autoimmunity. Although clonally expanded T cells have been identified in human autoimmune disease, the majority of memory T cells at sites of inflammation are not clonally expanded, suggesting a substantial contribution of "bystander" memory cells of unrelated antigenic specificity to chronic inflammation (*15*). The study of ElTanbouly *et al.* suggests that therapeutic targeting of VISTA may limit the generation of these bystander cells and could be a potential strategy for suppressing inflammation. ■

REFERENCES AND NOTES

1. M. A. ElTanbouly *et al.*, *Science* **367**, eaay0524 (2020).
2. K. Takada, S. C. Jameson, *Nat. Rev. Immunol.* **9**, 823 (2009).
3. P. Bretscher, M. Cohn, *Science* **169**, 1042 (1970).
4. K. J. Lafferty, J. Woolnough, *Immunol. Rev.* **35**, 231 (1977).
5. R. Medzhitov, P. Preston-Hurlburt, C. A. Janeway Jr., *Nature* **388**, 394 (1997).
6. J. Attanasio, E. J. Wherry, *Immunity* **44**, 1052 (2016).
7. Q. Zhang, D. A. A. Vignali, *Immunity* **44**, 1034 (2016).
8. L. Wang *et al.*, *J. Exp. Med.* **208**, 577 (2011).
9. L. Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 14846 (2014).
10. X. Han *et al.*, *Sci. Transl. Med.* **11**, eaax1159 (2019).
11. D. B. Flies *et al.*, *J. Clin. Invest.* **124**, 1966 (2014).
12. T. Kawabe *et al.*, *Sci. Immunol.* **2**, eaam9304 (2017).
13. X. Feng *et al.*, *Nat. Immunol.* **12**, 544 (2011).
14. R. J. Johnston *et al.*, *Nature* **574**, 565 (2019).
15. N. Saligrama *et al.*, *Nature* **572**, 481 (2019).

ACKNOWLEDGMENTS

C.B. is supported by a Wellcome Trust Fellowship; A.Y.R. is supported by the Howard Hughes Medical Institute, the Ludwig Center at Memorial Sloan Kettering, and the Parker Institute for Cancer Immunotherapy (PICI). This work was supported by NIH–National Cancer Institute (NCI) Cancer Center Support Grant (CCSG) P30CA008748.

10.1126/science.aaz7741

GRAPHIC: C. BICKEL/*SCIENCE*

Published by AAAS

EVOLUTIONARY BIOLOGY

High-resolution dating of Paleozoic fossils

Refined ages of marine fossils clarify the timing of diversification and extinction events

By **Peter Wagner**

Discussions about the quality of the fossil record frequently focus on completeness—that is, the proportion of species that existed for which scientists have fossil samples. An equally important aspect of quality is how finely the ages of fossils and durations of major evolutionary events can be resolved. Paleontologists and other biologists typically date fossils using the general ages of the chronostratigraphic units assigned to the rock strata yielding the fossils. For example, if a fossiliferous bed comes from the Rhuddanian Gasworks Sandstone, then the ages assigned to the fossils are usually 443.8 to 440.8 million years. On page 272 of this issue, Fan *et al.* (*1*) report on their use of constrained optimization (CONOP) (*2*) in biochronological analyses of fossil-bearing layers containing 11,000+ marine invertebrate species from 3000+ sections of sedimentary rocks, which constrain ages to a fraction of that range.

Their data, taken from the Geobiodiversity Database (GBDB) (*3*), span the Paleozoic through the Early Triassic. Although GBDB data are limited (largely) to China, the different continental plates that compose modern China fortuitously span multiple biogeographic realms throughout most of the Paleozoic. CONOP aligns fossil-bearing beds from different rock sections (e.g., spans of rocks from quarries, road cuts, and natural erosion on hills and mountains) by maximizing sets of species with first and last appearances in the same order in as many sections as possible. By integrating these analyses with radiometric dates and other geological data (*4*), the authors resolved the age of an average bed (and thus the average first and last appearance times of species) to within 26,000 years. This new level of dating specificity is similar to moving from a system in which all people who lived in the same century are considered to be contemporaries to one in which only people who lived during the same 6-month period are deemed to be contemporaries.

Armed with exact first and last appearance dates for thousands of species, the authors then developed what are, as they themselves stress, a set of fairly preliminary analyses of Paleozoic diversity. The authors deliberately eschew sophisticated analyses of origination and extinction dynamics or sampling dynamics used in other works with much coarser temporal resolution than that yielded by CONOP (*5*, *6*). This approach permits contrasting the basic patterns with those that would be implied had the authors chosen to "bin" these data into predefined chronostratigraphic analyses, as nearly all prior diversity studies have done.

"...constrained optimization results indicate rapid declines in richness of marine invertebrate species prior to the end-Ordovician and end-Permian extinctions."

In doing so, some noteworthy patterns emerge. For example, despite failing to correct for the Signor-Lipps effect (imperfect sampling that causes sudden extinctions to look gradual or rapid radiations to look protracted) (*7*), CONOP results indicate rapid declines in richness of marine invertebrate species prior to the end-Ordovician and end-Permian extinctions. Similarly, Fan *et al.* found the rebound after the end-Ordovician extinction to be much more rapid than prior studies suggest. Furthermore, CONOP reconstructions show patterns over hundreds of thousands of years that prior studies, which lumped together occurrences over millions of years, cannot hope to reveal.

Analysis of the Great Ordovician Biodiversification Event (GOBE) (*8*) offered a different type of unexpected result. Fan *et al.*'s study suggests that the GOBE began 20+ million years earlier than expected, during the late Cambrian. Given that late Cambrian sampling for many marine invertebrates is relatively poor compared to their fossil records earlier in the Cambrian and in the Ordovician (*9*), one would predict emergence of an opposite pattern: Any late Cambrian diversification should not have appeared until the Ordovician.

The CONOP analyses of GBDB data have implications for how organismal biologists and molecular phylogeneticists perform analyses of species, including fossils, in the future. Recent years have seen a proliferation of methods such as tip-data (*10*) and fossilized birth-death models (*11*), which include fossilized anatomical data and modern anatomical and molecular data. Accounting for uncertainty in the true ages of first and last appearances can strongly affect the results of such analyses (*12*). Current approaches for dealing with this uncertainty often assume that uncertainty is large; for example, if a species's first appearance is in the Rhuddanian age, then all dates within the Rhuddanian are equally probable first appearances (*13*). However, as the new study makes clear, biochronological analyses will make some dates within the Rhuddanian much more probable than others.

Of course, many extant organisms do not fossilize as well as shell-bearing marine invertebrates, and their oldest fossil representatives often are known from exceptional preservation in the lagerstätte sedimentary deposits (*14*). CONOP and related biochronology methods (*15*) require species from multiple beds in multiple sections to infer ages for fossil occurrences. However, fossiliferous beds preserving typical assemblages of shell-containing invertebrates usually occur above and below the lagerstätte. Constraining the ages of those beds will simultaneously constrain the ages of the lagerstätte. With respect to important evolutionary questions, the development of new datasets and methods of the sort used by Fan *et al.* will further evolutionary biology as a whole, not just paleontology. ■

Department of Earth and Atmospheric Sciences, and School of Biological Sciences, University of Nebraska Lincoln, Lincoln, NE 68588-0340, USA.
Email: peterjwagner@unl.edu

GRAPHIC: N. DESAI/*SCIENCE*

REFERENCES AND NOTES

1. J.-x. Fan *et al.*, *Science* **367**, 272 (2020).
2. P. M. Sadler, R. A. Cooper, in *High-Resolution Stratigraphic Approaches in Paleontology*, P. Harries, Ed. (Plenum, 2003), vol. 21, pp. 49–94.
3. GBDB, www.geobiodiversity.com.
4. P. M. Sadler, *Bull. Geosci.* **87**, 681 (2012).
5. M. Foote *et al.*, *Proc. Biol. Sci.* **285**, 20180122 (2018).
6. M. Foote *et al.*, *J. Geol. Soc. London* **176**, 1038 (2019).
7. P. W. Signor, J. H. Lipps, *Spec. Pap. Geol. Soc. Am.* **190**, 291 (1982).
8. T. Servais *et al.*, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **294**, 99 (2010).
9. Y.-D. Zhang *et al.*, *Palaeoworld* **28**, 1 (2019).
10. F. Ronquist *et al.*, *Syst. Biol.* **61**, 973 (2012).
11. T. A. Heath, J. P. Huelsenbeck, T. Stadler, *Proc. Natl. Acad. Sci. U.S.A.* **111**, E2957 (2014).
12. J. Barido-Sottani *et al.*, *Proc. Biol. Sci.* **286**, 20190685 (2019).
13. D. Silvestro *et al.*, *Paleobiology* **45**, 546 (2019).
14. D. T. Ksepka *et al.*, *Syst. Biol.* **64**, 853 (2015).
15. J. Alroy, *Paleobiology* **20**, 191 (1994).

10.1126/science.aba4348

Published by AAAS

HYPOTHESIS

Are noncommunicable diseases communicable?

Numerous noncommunicable diseases could have a transmissible microbial component

By **B. B Finlay[1,2]** *and* **CIFAR Humans and the Microbiome[2]**

The past century has seen a profound decrease in mortality rates across the world, accompanied by a marked shift from communicable diseases (caused by infectious microbes) to noncommunicable diseases (NCDs) such as cardiovascular diseases, cancer, and respiratory diseases. NCDs—defined as diseases that are not transmissible directly from one person to another—account for more than 70% (41 million) of all deaths globally (*1*). The definition of NCDs rules out microbial involvement and instead focuses on genetic, environmental, and lifestyle factors. Data increasingly show that the microbiota is dysbiotic (altered) in individuals with various NCDs. In animal models of NCDs, transplantation of dysbiotic microbiota into healthy animals results in disease, and microbiota composition is shaped by close contact with others. Therefore, we propose that some NCDs could have a microbial component and, if so, might be communicable via the microbiota.

Infectious diseases are caused by the transmission of pathogens between individuals. However, the extent to which microbial dispersal between humans contributes to NCDs remains unclear. The human microbiota consists of the various microbes (including bacteria, fungi, and viruses) living in and on the human body and has an important role in many physiological functions, including digestion, immune responses, and metabolism. Although microbes reside on many body sites, the majority are in the gut, with bacteria being the most studied. Several examples of NCD transmission exist by fecal microbiota transplant (FMT) into animal models, but how transmissible is the human microbiota? Cohabitants and spouses have more similar gut bacterial microbiota than genetically similar siblings living separately. Microbiota are transmissible within both family and social networks, and spousal relationships can be determined on the basis of gut bacterial analysis (*2*). Because families share diets and environments, their microbiota is expected to be similar. Thus, whether shared microbiota influence the transmissibility of NCDs is challenging to investigate, because uncoupling environment from microbiota is difficult.

Obesity is a leading risk factor for many NCDs, and there is increasing evidence that obesity has a microbial component. FMT from genetically predisposed or diet-induced obese animals to germ-free, lean animals causes significant weight gain (*3*), indicating that gut microbes are part of the etiology. The risk of postdieting weight regain in formerly obese mice is increased by a persistently altered gut microbiota, which is transferable to germ-free mice (*4*). Studies suggest that obesity may also be communicable in humans. In a social network study of 12,067 people over 30 years (*5*), having an obese friend was associated with a 57% higher chance of being obese, and there was a 40% higher chance of obesity if a sibling was obese. Moreover, a study of U.S. military families showed that being stationed in a county with high obesity rates was associated with an increased body mass index (BMI), whereas those stationed in counties with lower obesity rates had a lower BMI (*6*). These data are consistent with the idea that a socially transmissible component contributes to obesity, representing a shared environment, including diet and lifestyle, as well as microbiota. However, it is difficult to uncouple environment (diet, social habits) from microbiota composition, because they are intimately connected. Currently, microbial transmission of NCDs has only been demonstrated in controlled FMT experiments in genetically similar animal models with the same diets and environments.

Obesity is the highest risk factor for type 2 diabetes (T2D). Thus, the risk for developing T2D may also have a communicable component through the microbiota. Within a year of a T2D diagnosis, spouses have a higher chance of developing T2D, and this trend remains over 3 years after the initial diagnosis. In mice, T2D has a microbially transferable component, as demonstrated by FMT from mice with T2D into germ-free mice (*7*). Inflammatory bowel diseases (IBDs) are associated with characteristic dysbiotic microbiota, which can be transferred from diseased humans or mice to healthy animals along with the disease phenotype (*8*). Spouses of IBD patients have similar dysbiotic microbiota compositions and a higher rate of disease than accounted for by chance alone, although, like most infectious diseases, the "transmission" rate is not 100%. In India, the rate of ulcerative colitis (UC) is low, yet after moving from India to the United Kingdom, United States, and Canada, migrants have higher levels of UC. This change is attributed to "environmental factors," including diet and lifestyle, and the gut microbiota could be a contributing factor. Host genetic predisposition to IBD, and thus individual physiology, also plays a role, with more than 200 genetic loci linked to IBD (*9*). Many of these, such as *NOD2* (nucleotide-binding oligomerization domain–containing 2), are linked to immune functions that affect the gut microbiota composition, emphasizing the link between host genetics and the microbiota.

How can connections between transmissible microbiota and NCDs be tested? In 1890, Robert Koch published a set of postulates to determine whether a microbe was the cause of an infectious disease. Although exceptions exist, this set of "rules" has served well for establishing the causative agent of most infectious diseases. Applying a version of Koch's postulates that are adapted to NCDs could determine whether the collective microbiota can be considered an "infectious agent," which would support the hypothesis of communicable NCDs (see the figure). For these "microbiota-associated postulates," dysbiotic microbiota is considered the pathogen or causative agent and is defined as being different from the microbiota composition of unaffected individuals. The first postulate states that the microorganism should be present in those with disease. A strong correlation exists between a dysbiotic microbiota and many NCDs, including cardiovascular disease (CVD) and IBD. The second postulate states that the organism can be isolated from a diseased host and grown in pure culture. Collectively, dysbiotic microbiota can be harvested from feces, and many members grown. The third postulate states that the microbe, when inoculated into a healthy organism, should cause disease. FMT of dysbiotic microbiota from individuals with various NCDs into healthy animals results in disease, such as CVD, IBD, T2D, and many others. The final postulate states that the microorganism be isolated from the diseased host. This has been well documented for dysbiotic microbiota for many animal models of NCDs (*10*).

These modified postulates can be applied

[1]Michael Smith Laboratories and the Departments of Biochemistry and Molecular Biology, and Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada. [2]Fellows of the Canadian Institute for Advanced Research (CIFAR), Toronto, ON M5G 1M1, Canada. The list of authors and affliations is in the supplementary materials. Email: bfinlay@msl.ubc.ca

Published by AAAS

to CVD (heart attacks and strokes), the most prevalent NCD worldwide. There are strong correlations with the prevalence of particular gut microbiota that encode the enzyme choline trimethylamine (TMA)–lyase (CutC) that metabolizes phosphocholine and carnitine (from red meat) into TMA, which then undergoes hepatic oxidation into trimethylamine oxide (TMAO) (*11*). The concentration of TMAO in the blood is a strong predictor of CVD, with higher prevalence of disease associated with the presence of CutC-encoding gut microbes (*12*). Germ-free animals do not acquire CVD, even if on a choline-rich diet, and vegans and vegetarians have lower CVD rates than meat eaters (*13*). If CutC is inhibited in animal models, CVD does not occur. Moreover, human gut microbes encoding CutC can be transplanted into animals, leading to CVD phenotypes (*14*). Research examining spousal or community rates of CVD have so far only examined environmental effects (e.g., smoking, obesity, and alcohol). These may alter the gut microbiota composition, and so further research is warranted to examine whether microbial transmission is also involved.

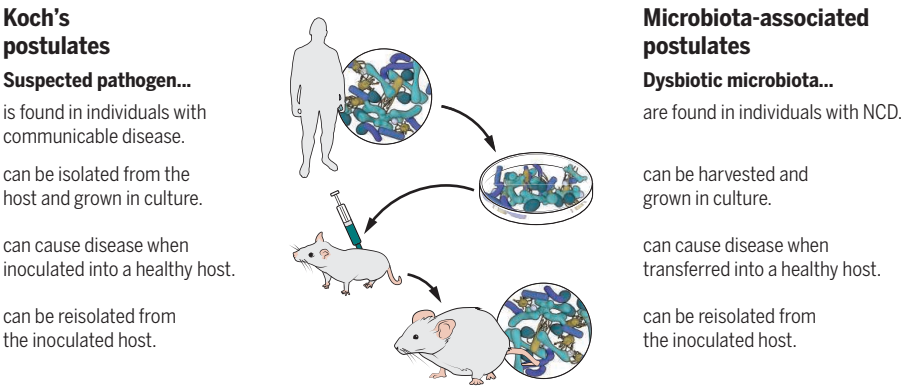
Applying modified Koch's postulates to CVD therefore reveals a strong correlation with a dysbiotic microbial composition (prevalance of bacteria encoding the CutC enzyme) and TMAO production and CVD, which addresses the first postulate. These organisms can be grown in the laboratory, thus satisfying the second postulate, and then transferred into healthy animals, which results in CVD (*14*), thereby satisfying the third postulate. These CutC-encoding microbes can also be isolated and cultured from diseased animals, satisfying the fourth postulate.

Similarly, these modified Koch's postulates can be applied to obesity (although not an NCD, it is the leading risk factor for many NCDs) and to IBD. In both cases, a characteristic dysbiotic microbiota can be introduced into animals, resulting in obesity or IBD. There are certainly caveats to this "proof" of causation and thus communicable NCDs. In particular, dysbiotic microbiota is not well defined and may differ in composition between individuals. Dysbiotic microbiota are associated to varying degrees with different NCDs—for example, they play a smaller role in cancer than in CVD. These analyses raise the hypothesis that transmissible dysbiotic microbiota contribute to NCDs in humans, but uncoupling this from environmental components and genetic predisposition will require substantial additional research.

These observations suggest that the microbiota could be a causal and transmissible element in certain diseases that have been traditionally classified as NCDs. It is hoped that this hypothesis stimulates additional discussion and research, including studies that define environmental effects on the microbiota, identifying microbial members that constitute a dysbiotic transmissible microbiota that confers disease, and further delineating the extent of the contribution of the microbiota to NCDs. Notably, transmissible microbiota, especially early in life, may also have a protective role against NCDs, including asthma, allergies, and obesity; these protective microbes can also be experimentally transmitted to animal models (*15*). Additionally, only gut bacteria have been considered in this discussion, yet viruses and fungi may also contribute to NCDs, as well as microbiota at other body sites such as the skin and oral cavity. As the potential role of transmissible microbiota in NCDs becomes better defined, it will provide new opportunities to address these complex diseases. ■

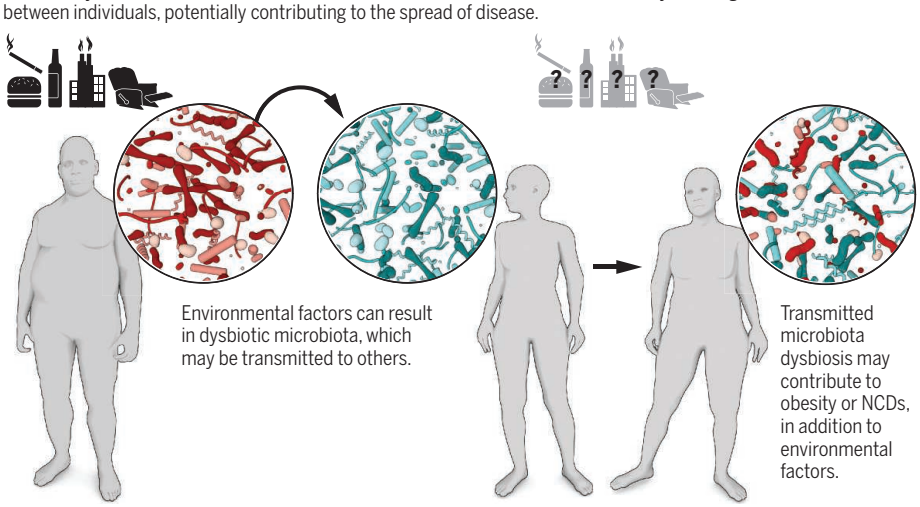
Transmission of dysbiotic microbiota

It is proposed that dysbiotic microbiota, particularly in the gut, can be transmitted to other individuals, in turn altering their microbiota. This may contribute to the spread of noncommunicable diseases (NCDs), including obesity, a key risk factor for many NCDs, as well as possibly cardiovascular disease and inflammatory bowel diseases.



Microbial transmission in NCDs?

The microbiota of an individual is affected by environmental factors such as diet, smoking, alcohol intake, and exercise. Dysbiotic microbiota can influence NCDs or their risk factors, such as obesity, and might be transmitted between individuals, potentially contributing to the spread of disease.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

REFERENCES AND NOTES

1. WHO, https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases.
2. I. L. Brito *et al.*, *Nat. Microbiol.* **4**, 964 (2019).
3. P. J. Turnbaugh *et al.*, *Nature* **444**, 1027 (2006).
4. C. A. Thaiss *et al.*, *Nature* **540**, 544 (2016).
5. N. A. Christakis, J. H. Fowler, *N. Engl. J. Med.* **357**, 370 (2007).
6. A. Datar, N. Nicosia, *JAMA Pediatr.* **172**, 239 (2018).
7. C. A. Thaiss *et al.*, *Cell* **159**, 514 (2014).
8. E. Elinav *et al.*, *Cell* **145**, 745 (2011).
9. A. N. Ananthakrishnan, *Nat. Rev. Gastro. Hep.* **12**, 205 (2015).
10. P. Vonaesch, M. Anderson, P. J. Sansonetti, *FEMS Microbiol. Rev.* **42**, 273 (2018).
11. W. Zhu *et al.*, *Cell* **165**, 111 (2016).
12. Z. Wang *et al.*, *Nature* **472**, 57 (2011).
13. H. Kahleova, S. Levin, N. D. Barnard, *Prog. Cardiovasc. Dis.* **61**, 54 (2018).
14. J. C. Gregory *et al.*, *J. Biol. Chem.* **290**, 5647 (2015).
15. M. C. Arrieta *et al.*, *Sci. Transl. Med.* **7**, 307ra152 (2015).

ACKNOWLEDGMENTS

We acknowledge funding from the Canadian Institute for Advanced Research (CIFAR).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/250/suppl/DC1

10.1126/science.aaz3834

Published by AAAS

POLICY FORUM

OCEAN GOVERNANCE

Mobile protected areas for biodiversity on the high seas

Protecting mobile marine species and habitats under climate change will require innovative and dynamic tools

By **Sara M. Maxwell[1], Kristina M. Gjerde[2], Melinda G. Conners[3], Larry B. Crowder[4]**

A new agreement is being negotiated under the 1982 United Nations Convention on the Law of the Sea (UNCLOS) to provide legally binding mechanisms to protect the marine environment and to conserve and ensure the sustainable use of marine biodiversity on the high seas (international waters in areas beyond national jurisdiction) (*1*). One of the suggested objectives in the current draft text is to "apply an approach that builds ecosystem resilience to the adverse effects of climate change" when applying area-based management tools (ABMTs), including marine protected areas (MPAs). Yet even though climate change is resulting in shifts in species' ranges (*2*) and in the behavior of the human users of mobile, commercially valuable species (*3*), protection of highly mobile species and the dynamic habitats on which they depend is not currently a focus of negotiations. With the final language to be determined as early as 2020 (*1*), we urge negotiators to include new dynamic management tools, including mobile MPAs (mMPAs), whose boundaries shift across space and time, that could help to safeguard marine life and build ecosystem resilience by protecting dynamic habitats as well as migratory marine species in a changing ocean.

Large MPAs (>10,000 km^2), which have been increasing globally in number, will be an important addition to the toolbox for protecting habitat and building resilience of many high-seas species and ecosystems (*4*). However, key habitats of highly mobile species are unlikely to be protected by static boundaries alone. Some migratory species range entire ocean basins (*5*); in addition, animal distributions may change considerably, either over short time scales due to seasonal or interannual climate variability, or over the long term as a result of climate change (*2*). In the past decade, dynamic spatial management of ocean resources, which responds to changes in marine systems or resources on short time scales (days to seasons), has been a critical tool for managing fisheries in national waters. It has also been applied to reduce biodiversity impacts, such as through the closure of fishing areas based on oceanographic conditions that correlate with high bycatch, or reduced vessel speeds when whales are detected in shipping lanes (*6*).

For many wide-ranging species, dynamic ABMTs without static boundaries may be the only practical option for sufficient protection, although few UN negotiators are familiar with the advances in science and technology that make dynamic ABMTs possible for the high seas. By recognizing, defining, and enabling flexible dynamic area-based approaches, UN negotiators have a unique opportunity to complement the efforts of international sectoral organizations [e.g., regional fishery management organizations (RFMOs)] to achieve sustainable ecosystem-based management (e.g., by avoiding bycatch of vulnerable migratory species), as well as increasing the effectiveness of traditional MPAs during this time of unprecedented challenges attributable to climate change (*2*).

MOBILE PROTECTION ON THE HIGH SEAS

On the high seas, dynamic ABMTs occur infrequently and under the management of individual countries, and only for fisheries (countries cannot control the high-seas area itself but can control the activities of their flagged vessels). In the Australian multispecies longline fishery, mandated fishing zones that extend into the high seas are determined using forecasted southern bluefin tuna (*Thunnus maccoyii*) habitat in near-real time to not exceed quotas (*7*).

PHOTO: TUI DE ROY/MINDEN PICTURES

TurtleWatch is a voluntary program updated weekly and applied by the United States to its longline vessels in the North Pacific to reduce sea turtle bycatch based on turtles' sea surface temperature preferences (*8*). These examples highlight that most of the technological limitations to the application of dynamic ABMTs that were recognized nearly 20 years ago (*9*) have been overcome through advances in animal tracking, satellite imagery, computing capacity, and communication. These have led to the ability to tailor management efforts by determining where species or dynamic habitats such as fronts occur (*6*) or by detecting mobile human users (*10*). Dynamic ABMTs could become a vital conservation tool for the high seas if legal, political, and scientific obstacles can be addressed through the UNCLOS implementing agreement for marine biodiversity.

Including mMPAs—dynamic management that focuses on the comprehensive conservation objectives inherent in MPAs—as a potential tool in the ABMT toolbox will be critical for moving to an approach that enables comprehensive protection of marine biodiversity as species, habitats, and ecological communities shift in a changing environment. Some studies have suggested that current MPAs, including those

[1]School of Interdisciplinary Arts and Sciences, University of Washington Bothell, Bothell, WA, USA. [2]IUCN Global Marine and Polar Programme, World Commission on Protected Areas, Cambridge, MA, USA. [3]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA. [4]Hopkins Marine Station, Stanford University, Pacific Grove, CA, USA. Email: smmax@uw.edu

Published by AAAS

Campbell albatross (*Thalassarche impavida*) are found on Bull Rock, Cape Colony, Campbell Island, New Zealand.

on the high seas, will continue to protect biodiversity under climate change scenarios (*4*); others suggest that current MPAs will be ineffective, particularly in tropical and temperate regions (*11*), reflecting substantial uncertainty around the nuances of how species are responding to climate change. As species and habitats shift across space and time, mMPAs could be used to protect dynamic oceanographic habitats critical for ecosystem function, such as fronts, currents, or eddies, or to protect individual species or groups of species. mMPAs can further serve to protect connectivity corridors between static MPAs and thereby help to "future-proof" such MPAs against shifts due to climate change by offering protection when key species or habitats shift outside static boundaries. Dynamic habitats have already been described on the high seas through the Convention on Biological Diversity's (CBD) Ecologically and Biologically Significant Area process (*12*) (see the figure and the supplementary table). Initiatives such as the Migratory Connectivity Project (MiCO) of marine species can be built upon to track the movement of species across their ranges and jurisdictional boundaries (*13*). Like traditional MPAs, mMPAs could be managed to protect species and habitats from multiple threats, including those that involve multiple sectors (e.g., industrial fishing, shipping, seismic surveys) as well as those from diffuse or multiple sources, such as ocean noise. Within mMPA boundaries, human activities would be restricted, similar to static MPAs, but these boundaries (and the applicable measures inside them) would move, tailored to the movement of the habitat or species being protected. Protecting species that move has been a tenet of conservation biology for decades, but having protections follow them in near-real time has not been possible because we lacked the ability to, for example, communicate to users where the boundaries were as they moved. These hurdles, however, have been overcome in the digital age.

IMPLEMENTATION WITHIN UNCLOS

We recommend that dynamic ABMTs, including mMPAs, should be recognized as a potential tool within the UNCLOS implementing agreement by defining ABMTs to clearly include (or at least not exclude) spatially or temporally variable measures and include as an objective the protection of ecosystems, natural habitats, and populations of migratory species throughout their range. To do this, we recommend the following: (i) A Conference of the Parties (COP) to the implementing agreement should be empowered to establish ABMTs, including mMPAs, and call on states parties to adopt relevant conservation measures (applicable to their flagged vessels and nationals). (ii) The COP should also be able to recommend that sectoral management organizations (e.g., RFMOs, International Maritime Organization) adopt dynamic ABMTs for specific species or habitats based on globally agreed conservation priorities, criteria, and guidelines adopted under the agreement. (iii) A scientific expert body should have a key role in reviewing proposals and advising on implementation.

Potential sector-based dynamic ABMT measures might include changes in ship-routing measures on short time scales based on the distribution of whales, discharge limitations in areas identified as key foraging grounds of sensitive species, and intermittent gear restrictions to avoid bycatch. Such measures could be implemented directly by states parties, as is the case with the Australian longline fishery and Turtle-Watch in the United States (*7*, *8*), as well as recommended for adoption more widely by the relevant international sectoral organization. MPAs, including mMPAs, could include more comprehensive conservation measures that address multiple threats across sectors and be informed by a targeted management, monitoring, and research plan, as with traditional MPAs. Including a specific obligation to adopt dynamic measures in a global agreement could "institutionalize" dynamic ABMTs at the international level, while also advancing the currently inconsistent implementation of several existing obligations: to "protect and preserve rare or fragile ecosystems as well as the habitat of depleted, threatened or endangered species and other forms of marine life" under UNCLOS; to establish effective protected areas and take other measures to "promote the protection of ecosystems, natural habitats, and the maintenance of viable populations of species in natural surroundings" under the CBD; and to adopt measures to conserve migratory species and their habitats listed in the Convention on Migratory Species and its sister agreements (*14*).

Dynamic ABMT boundaries could be defined in a number of ways. These might include demarcating boundaries by explicit environmental characteristics, such as sea surface temperature bands, rather than by static latitude and longitude coordinates; by determining the presence of specific species by visual or acoustic detection; or by predicting habitats or species occupancy through modeling or forecasting (*6*). Such areas have distinct geographic boundaries,

Published by AAAS

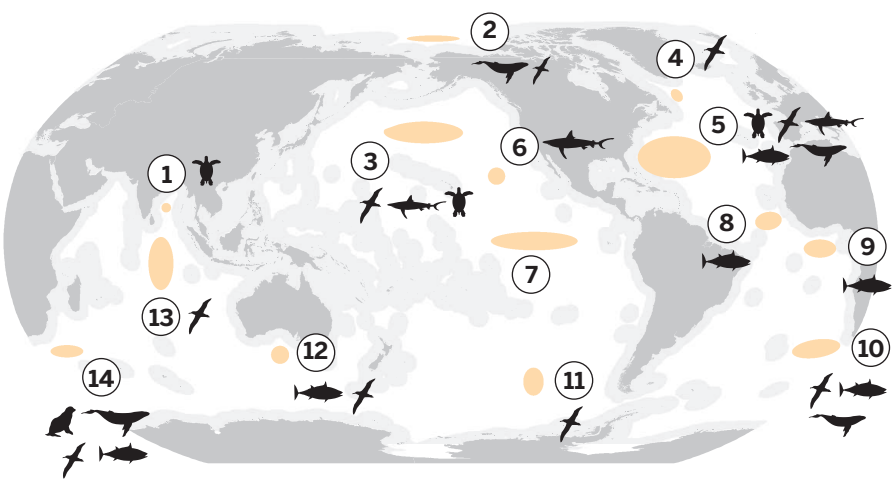
although they move in space and time; for mMPAs, the criteria or conditions that define where boundaries are placed could be persistent, allowing for mMPAs to meet the International Union for Conservation of Nature (IUCN) definition of an MPA. Notably, management measures may only be necessary during certain times (e.g., during a species' breeding season) and in certain places, potentially resulting in a smaller footprint of restrictive management of human activities, as has been shown for dynamic management more broadly (*15*). Managing for direct impacts on marine species and habitats can improve species resiliency by reducing the direct threats that lead to population decline, thereby providing a buffer against indirect impacts from climate change such as changes in prey distribution.

Dynamic ABMTs cannot directly address a number of issues, such as movements of species or habitats of concern across political boundaries such as Exclusive Economic Zones, potential for lack of monitoring and enforcement (including illegal, unreported, and unregulated industries), or ineffective management. Nor will they displace the need for more traditional, stationary ABMTs including static MPAs to protect fixed habitat features such as seamounts, less mobile species, or areas of cultural significance. Conflicts may arise around communicating boundaries as they shift over time, but all management relies on effective communication with human users, and existing dynamic management applications have applied technologies successfully, including sophisticated websites, smartphone and tablet applications, and simple email and cell phone communications (*6*). Additionally, in offshore pelagic environments such as the high seas, safe operation requires most vessels to be of a length that legally requires automatic ship identification systems (AISs). Now that AIS data are widely available, we have the capacity for effective near–real time monitoring, and potentially enforcement, without the need for at-sea enforcement missions, although some legal frameworks may need to be enhanced to require AIS use by all vessels on the high seas (*10*).

In domestic waters, dynamic management is increasingly applied, and this experience can underpin its application on the high seas (*6*). Thus, the new treaty being considered by UN negotiators presents an opportunity to embrace global commitments under the CBD, the UN Sustainable Development Goals, and the upcoming post-2020 global biodiversity framework by advancing beyond the traditional static geographic constraints on ABMTs and MPAs. By enabling dynamic ABMTs, including mMPAs, the global community can together build resilience and maintain, conserve, and restore increasingly vulnerable migratory marine species in the context of a changing ocean. ■

Dynamic habitats on the high seas

Protection of species or habitats in such areas may benefit from dynamic area–based management tools or mobile marine protected areas. Areas are identified from the Convention on Biological Diversity's Ecologically or Biologically Significant Marine Areas (www.cbd.int/ebsa/).



1 Olive Ridley sea turtle migration corridor
Mass migration corridor for globally significant population of vulnerable sea turtles.

2 Deep Arctic Marginal Ice Zone and seasonal sea-ice cover
Dynamic habitat critical to many species; ice structure altering rapidly under climate change.

3 North Pacific Transition Zone
Productive migratory species habitat, shifts across seasons and years; shifting north under climate change.

4 Labrador Sea Seabird Foraging Zone
Productive wintering and foraging area for more than 40 million seabirds annually.

5 Sargasso Sea
Highly dynamic, unique, and biodiverse. Critical habitat for species across taxa.

6 Northeast Pacific white shark aggregation area
Seasonal aggregation area for white sharks, potentially for mating or foraging.

7 Equatorial High-Productivity Zone
Unique and highly productive, influenced by El Niño–Southern Oscillation events. Susceptible to climate change influence.

8 Canary-Guinea Current Convergence Zone
Strong upwelling area, supports commercially important fishes.

9 Equatorial production area
High-productivity migratory, spawning, and nursery habitat for commercially important fishes.

10 Subtropical Convergence Zone
High-productivity oceanographic feature that supports endangered seabirds and other species.

11 Southeast Pacific Rise Grey Petrel Feeding Area
Critical foraging area for Near Threatened seabird species from October to February.

12 South of Great Australian Bight
Critical foraging albatross and migrating tuna habitat.

13 Central Indian Ocean Basin
Important seabird foraging area, heavily influenced by seasonal productivity during austral winter.

14 Agulhas Front
Highest productivity in Indian Ocean; supports diversity of seabirds, mammals, and tunas.

GRAPHIC: N. DESAI/*SCIENCE*

REFERENCES AND NOTES

1. United Nations, *International Legally Binding Instrument Under the United Nations Convention on the Law of the Sea on the Conservation and Sustainable Use of Marine Biological Diversity of Areas Beyond National Jurisdiction*, Resolution 72/249 (2017), p. 4.
2. E. S. Poloczanska *et al.*, *Nat. Clim. Chang.* **3**, 919 (2013).
3. W. W. L. Cheung *et al.*, *Glob. Change Biol.* **16**, 24 (2010).
4. T. E. Davies, S. M. Maxwell, K. Kaschner, C. Garilao, N. C. Ban, *Nat. Sci. Rep.* **7**, 9569 (2017).
5. A.-L. Harrison *et al.*, *Nat. Ecol. Evol.* **2**, 1571 (2018).
6. S. M. Maxwell *et al.*, *Mar. Policy* **58**, 42 (2015).
7. A. J. Hobday, J. R. Hartog, C. M. Spillman, O. Alves, *Can. J. Fish. Aquat. Sci.* **68**, 898 (2011).
8. E. A. Howell *et al.*, *Fish. Oceanogr.* **24**, 57 (2015).
9. K. D. Hyrenbach, K. A. Forney, P. K. Dayton, *Aquat. Conserv.* **10**, 437 (2000).
10. D. J. McCauley *et al.*, *Science* **351**, 1148 (2016).
11. J. F. Bruno *et al.*, *Nat. Clim. Chang.* **8**, 499 (2018).
12. D. C. Dunn *et al.*, *Mar. Policy* **49**, 137 (2014).
13. D. C. Dunn *et al.*, *Proc. R. Soc. B* **286**, 20191472 (2019).
14. K. M. Gjerde, N. A. Clark, H. R. Harden-Davies, *Ocean Yearb. Online* **33**, 1 (2019).
15. D. C. Dunn, S. M. Maxwell, A. M. Boustany, P. N. Halpin, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 668 (2016).

ACKNOWLEDGMENTS

Supported by a fellowship in ocean sciences from the Alfred P. Sloan Foundation, a grant from the Initiatives to Develop Interdisciplinary Scholarship and Collaboration (I-DISCO) program from the School of Interdisciplinary Arts and Sciences at University of Washington Bothell, and the Helen Riaboff Whiteley Center (S.M.M.); the Gallifrey Foundation (K.M.G.); a grant from the Pew Charitable Trusts (M.G.C.); and funds administered by the Woods Institute for the Environment (L.B.C.).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/252/suppl/DC1

10.1126/science.aaz9327

Published by AAAS

DATA SCIENCE

Patterns, statistics, and the stars

Highlights from the history of astrology reveal the origins of our quest to find meaning in data

By **Steven Vanden Broecke**

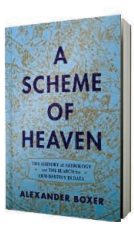
Astronomers and astrologers alike have relied on the astrolabe to plot schemes of the heavens.

PHOTO: G. DAGLI ORTI/DE AGOSTINI PICTURE LIBRARY/GETTY IMAGES

Alexander Boxer, a professional data scientist, knows a thing or two about distilling patterns from big data. Surrounded by constant, endless streams of information, humans are pattern-matching animals, and astrology, he claims, "is the universe's grandest pattern-matching game."

Boxer's new book, *A Scheme of Heaven*, is an introduction to astrology that offers a handsome primer of the structures and concepts with which astrologers have tried to make our world of data manageable. He constructs his story as a historical narrative, picking out specific historical events to reveal the degree to which local religious, political, social, and cultural contexts shaped astrology's pattern-matching game.

The book also exposes readers to the rigor of statistical analysis. Here, Boxer applies his knowledge of statistics to some of the most enduring and fascinating patterns that astrology educed from its constant comparisons between heavenly and terrestrial events. This combination of topics is usually the preserve of critics, who like to mobilize analyses of astrology's conceptual apparatus, history, and statistical soundness to demonstrate the art's vacuity. *A Scheme of Heaven* is different in that it seeks to offer a kind word for the endeavor.



A Scheme of Heaven: The History of Astrology and the Search for Our Destiny in Data
Alexander Boxer
Norton, 2020. 349 pp.

Astrology was the first art to capitalize on "the powerful storytelling possibilities inherent in numerical data," writes Boxer. In that regard, it was the predecessor of all modern big data disciplines. To be clear, Boxer is not a closet believer. He never argues, for example, for the reinstatement of the practice into the academy. (It was ousted in the second quarter of the 17th century.) His narrative takes a far more interesting tack.

Refreshingly, *A Scheme of Heaven* does not offer a comprehensive history of astrology but guides the reader along particularly important episodes in the practice's story, taking cues from specific historical events, characters, and objects. The resulting gallery tour takes us from ancient Egyptian pyramids and coffin lids to Mesopotamian cuneiform tablets. We learn of the cosmic models and analog calculators that Archimedes' workshop might have contained, and astrological intelligence policy in 1st-century imperial Rome (and its relation to the philosophical schools of the time) becomes more vivid than ever. Descriptions of the ambitious programs to systematize and explain human diversity through astrology in 2nd-century Egypt and to turn astrology into a comprehensive theory of human history in 8th-century Baghdad amply convey the historical fascination with astrological patternmaking. Late medieval Italy, on the other hand, becomes the backdrop for exploring astrology as a science of practical algorithms that would not be out of place in the modern world of stock forecasts, whereas the Age of Aquarius sees astrology reworked into an art of spiritual counseling, a shift that Boxer traces to the theosophical movements of the 1870s.

A Scheme of Heaven—like all good history writing—turns its subject into a mirror. (In the words of the Roman poet Horatius, "the story is told about you.") Statistics, Boxer shows, not only debunk astrology's claims, they confirm that some of our most private behavior happens in step with cosmic rhythms today. History not only documents a distant past, it shows how intimately some of our most prestigious scientific traditions really are—as Johannes Kepler argued—the children of this foolish daughter. And like astrology, the patterns that data science reveals turn out to hinge on far more interpretation than we might like. Boxer points out, for example, how the contemporary combination of big data with machine-learning algorithms is rapidly creating a rift between empirical forecasting models and causal understanding—exactly the kind of rift that has often been invoked to criticize astrology.

In the closing pages of this excellent book, Boxer candidly voices his fear that his fellow scientists will misunderstand his efforts as an endorsement of astrology. From my own experience, I suspect academia is more neutral toward such enterprises than it once was. Recognizing the historical importance of astrology certainly makes for a far messier account of early scientific practices. Done correctly, however, these messier pictures of science only reinforce our appreciation of the amount of work that goes into it—and of the vulnerability of human wonder propelling it. ■

The reviewer is at the Department of History, Ghent University, 9000 Ghent, Belgium. Email: steven.vandenbroecke@ugent.be

10.1126/science.aaz9644

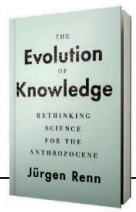
Published by AAAS

HISTORY OF SCIENCE

Knowledge in the Anthropocene

To save society, science should embrace ethical projects

The Evolution of Knowledge: Rethinking Science for the Anthropocene
Jürgen Renn
Princeton University Press, 2020. 579 pp.



By **Deborah R. Coen**

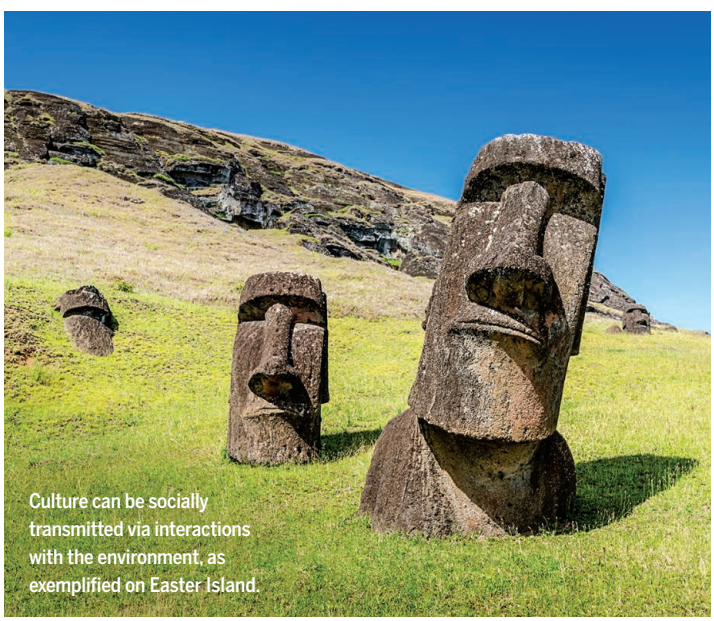
Can science save humanity? In the face of runaway climate change and massive species extinction, some say that we already know all we need to know to fix these problems: Further research is a distraction and what we need now is action. Others anticipate a feat of technical ingenuity that will catapult us out of our current crisis. In their view, the "Anthropos" of the Anthropocene is well on the way to mastering the Earth system. A middle ground between these positions hardly seems to exist.

Into this fray comes Jürgen Renn, a polymath historian of science, whose research has careened from ancient Chinese mechanics to Einstein's theory of gravitation to contemporary energy systems. What links these projects is Renn's fascination with the dynamics by which science changes and changes hands. His new tour de force, *The Evolution of Knowledge*, addresses all those concerned with science's fate.

Renn argues that the crisis of the Anthropocene is indeed a problem of knowledge, but he sharply distinguishes himself from those hunting for a technical fix. Knowledge, for Renn, is a broad and varied concept, with crucial experiential and ethical dimensions. Modern science, although uniquely efficacious, is just one facet of the evolution of human knowledge. Understanding this evolutionary process, he argues, is the key to reorienting science for the Anthropocene.

Taking a cue from biologists, Renn thinks of knowledge as an adaptive system, one that gradually transforms the material and cultural conditions of its own existence. Knowledge evolves through a slow, piecemeal process akin to ecological succession. If we want to guide the future evolution of science, we had better get familiar with its past.

Renn's history of knowledge is alternately triumphant and tragic. On the bright side, he traces the development of techniques of representation that have increasingly allowed for reflective engagement with knowledge-making. His foremost example is the internet. With the proper oversight, he proposes, the internet has the potential to "optimize the current knowledge economy toward a global coproduction of knowledge," building new connections between local stores of knowledge and new opportunities to put knowledge into action. What's more, science has evolved unprecedented power to reshape the "cultural abstractions" that govern human behavior—as, for instance, in the idea of an ecological footprint.



Culture can be socially transmitted via interactions with the environment, as exemplified on Easter Island.

Yet this history of modern science is also a story of loss. Market forces have distorted science's ambitions, diverting it from the humanitarian aims of earlier centuries. The globalization of science has suppressed local forms of knowledge and alienated nonexpert populations. Renn thus returns to the past in part to remind readers of the value of natural knowledge produced in other eras, by a panoply of human cultures, even by nonliterate societies.

One example of the insights to be drawn from Renn's historical case studies can be found in his retelling of the encounter between Jesuit and Chinese astronomers in the early 17th century. The Chinese adopted Copernican astronomy for the purpose of reforming official calendars, but they had no intention of allowing it to mingle with their religious views. Renn observes that this episode juxtaposes a remarkably stable intellectual tradition, that of the Ming dynasty, with one undergoing rapid change, as Europeans jettisoned medieval scholasticism in favor of rationalism and empiricism. The critical difference, he argues, was the tight relationship between natural knowledge and religion in the European scholastic tradition.

That observation brings us to one of Renn's most provocative proposals. He argues that modern science, rather than striving to be value-free, should embrace ethical projects of the sort usually associated with religion. Fully aware of the atrocities that could result from turning science into a religion, he nonetheless proposes that we "seek out the eschatological dimensions of science itself and cultivate its role as a guide in a fragile world whose future depends on it."

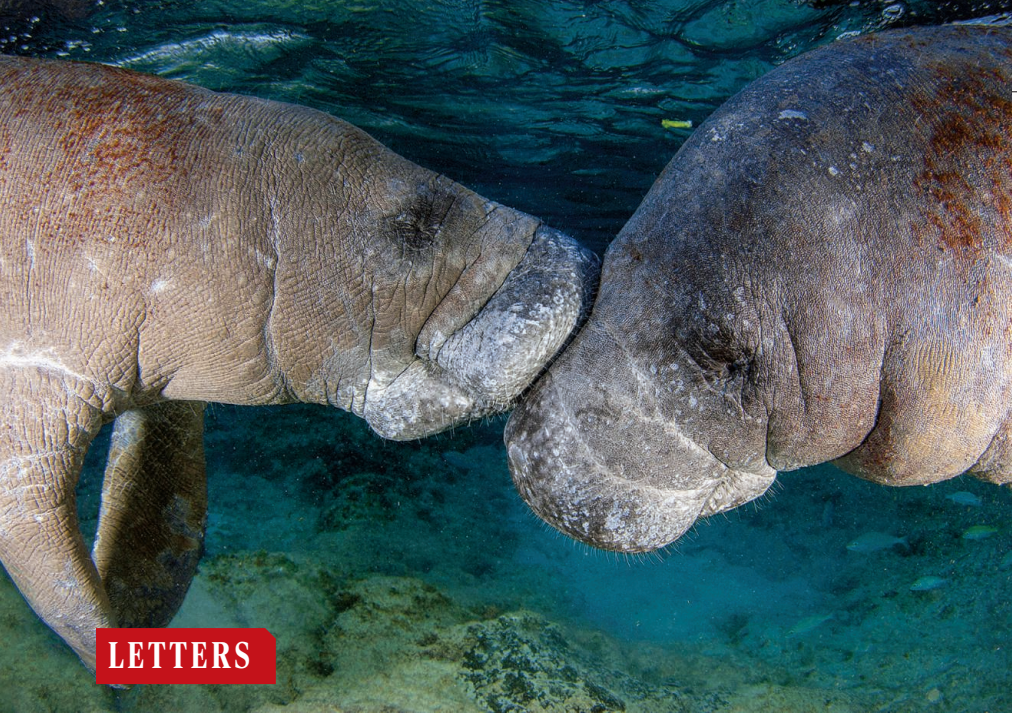
Ours is hardly the first era to witness calls for a wholesale reform of natural knowledge. From Francis Bacon in the 17th century to Vannevar Bush in the 20th, modern science has had its fair share of visionary reformers. In this respect, Renn might best be compared to the philosopher Edmund Husserl (1859–1938). In the 1930s, at a moment of existential crisis comparable to today's, Husserl likewise sought to reorient science around shared human experiences and common human needs. Yet Husserl, a notoriously opaque writer, had little hope of communicating his message to the scientific community. With this lucid and accessible book, Renn stands a far greater chance of success. ■

The reviewer is chair of the Program in the History of Science and Medicine, Yale University, New Haven, CT 06520, USA, and the author of *Climate in Motion: Science, Empire, and the Problem of Scale* (University of Chicago Press, 2018). Email: deborah.coen@yale.edu

10.1126/science.aba1244

PHOTO: ISTOCK.COM/MLENNY

Published by AAAS



LETTERS

PHOTO: RGB VENTURES/SUPERSTOCK/ALAMY STOCK PHOTO

Manatees in Mexico's waters face a range of threats, and their populations are declining.

Edited by **Jennifer Sills**

Losing Mexican manatees

Over the past 2 years, the Antillean manatee (*Trichechus manatus*) in Tabasco, Mexico, has shown an alarming increase in mortality. In 2018, the Mexican government reported 48 manatee deaths attributed to algal blooms (*1*). By the end of June 2019, another 13 manatees had been found dead (*2*). It is estimated that a few hundred manatees remain throughout the Mexican territory (*3*, *4*), but population counts have not been updated since 1999 (*5*). Continued threats put these manatees at risk of local extinction.

Despite the manatees' classification as endangered by Mexico (*6*) and vulnerable by the International Union for Conservation of Nature (*7*), Mexico has yet to invest the economic resources required to save them. Their habitat is exposed to contamination by agrochemicals, byproducts of oil industry, and urban waste such as pesticides, hydrocarbons, and toxic metals (*1*). Initially, the deaths in 2018 were attributed to toxic metals (*8*), but given that manatees can accumulate more toxic metals than other mammals (*9*), uncertainty about the cause of death remains.

Manatees are important to Mexico's culture, ecology, and tourism (*10*). The species requires urgent financial and technical support, starting with short- and medium-term diagnostics and monitoring studies to clarify the causes of mortality. Rescuing the manatee will require joint efforts by national and international private foundations and nongovernmental organizations. Mexican financial resources are limited; government investment in conservation of natural resources and science in general, combined, has not exceeded 0.55% of GDP in recent years (*11*).

The rescue of manatees in Tabasco has become urgent because the manatee is one of only four extant species of sirenids, along with the dugong (*Dugong dugon*), which is also threatened (*12*). Mexico, and international conservation allies, must prioritize manatee protection by addressing the effects of human activities and climate change on their habitat.

Gabriel Núñez-Nogueira* and Melina Uribe-López
Hydrobiology and Aquatic Pollution Laboratory, División Académica de Ciencias Biológicas, Universidad Juárez Autónoma de Tabasco, Entronque a Bosques de Saloya, 86150 Villahermosa, Tabasco, Mexico.
*Corresponding author.
Email: gabriel.nunez@ujat.mx

REFERENCES AND NOTES

1. Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT), Procuraduria Federal de Proteccion al Ambiente, "Reporte final: Atención de contingencia de manatíes en Tabasco" (2018); www.gob.mx/cms/uploads/attachment/file/410558/Reporte_mortandad_de_manaties_act__05_nov_18.pdf [in Spanish].
2. F. Bautista Villegas, "Muere otro manatí en Centla," XEVT Noticias (2019); www.xevt.com/verpagina.php?id=72928 [in Spanish].
3. SEMARNAT, Comisión Nacional de Áreas Naturales Protegidas, "Programa de acción para la conservación de la especie: Manatí (*Trichechus manatus manatus*)," L. D. Olivera Gómez, A. Ortega-Argueta, B. Morales Vela, L.C. Colmenero Rolón, Eds. (2010) [in Spanish].
4. M. J. Pola Tellechea, "Manatí, en peligro de extinción en Tabasco" (2018); www.diariopresente.mx/amp/villahermosa/manati-en-peligro-de-extincion-en-tabasco/188081 [in Spanish].
5. J. Morales-Vela *et al.*, "Informe final del Proyecto H164" (1999); www.conabio.gob.mx/institucion/proyectos/resultados/InfH164.pdf [in Spanish].
6. "Manatí del Caribe (*Trichechus manatus*)," NaturaLista (2000); www.naturalista.mx/taxa/46316-Trichechus-manatus [in Spanish].
7. C. J. Deutsch, C. Self-Sullivan, A. Mignucci-Giannoni, "*Trichechus manatus*" (The IUCN Red List of Threatened Species, 2008).
8. V. Santiago, "Mueren manatíes y otros animales en Tabasco por supuesta contaminación de metales" (2018); https://aristeguinoticias.com/1008/mexico/mueren-manaties-y-otros-animales-en-tabasco-por-supuesta-contaminacion-de-metales/ [in Spanish].
9. G. Núñez-Nogueira *et al.*, *Int. J. Environ. Res. Public Health* **16**, 404 (2019).
10. L. D. Olivera Gómez, "¿Por qué debemos conservar al manatí?," *Diàlogos* **40**, 12 (2012); www.ccytet.gob.mx/Docs/ccytet03/Dialogos/dialogos40.pdf [in Spanish].
11. V. Guadarrama, *INCYTU* **011**, 1 (2018) [in Spanish].
12. B. Li *et al.*, *Science* **365**, 552 (2019).

10.1126/science.aba2925

Support Austria's glyphosate ban

The herbicide glyphosate is the most used agrochemical herbicide weedkiller worldwide (*1*). Since 1974, more than 8.6 billion kilograms of glyphosate have been sprayed on crop fields (*2*). It was for decades thought to be a harmless alternative to legacy pesticides such as the banned DDT and parathion, which kill insects but also harm humans (*3*). However, new evidence shows that glyphosate causes a cascade of neuro-endocrine disruption to the development, physiology, and behavior of honeybees (*4*) and is thereby adding to the ongoing negative effects from neonicotinoids, which have led to the deaths of pollinators and songbirds (*5*, *6*). Moreover, some evidence has indicated that glyphosate could promote cancer in humans (*1*). If true, these compounds could pose a risk to human consumers as well. These effects are reminiscent of the events more than 50 years ago, when DDT caused substantial losses in biodiversity and ecosystem services (*7*). We must work to prevent history from repeating itself.

In 2017, the European Union reapproved its use of glyphosate for another 5-year period (*1*). Likewise, the U.S. Environmental Protection Agency states that glyphosate is below the levels of concern and has continued their use (*1*). Austria is the only EU member country that has passed a total ban of the herbicide. Although the Austrian parliament has voted to implement the ban in January (*8*), the EU Commission may try to veto it. The European Union and the United States should follow Austria's example and enact a total ban of glyphosate use, just like the international limitations that are currently underway to ban neonicotinoids (*9*, *10*). Such a ban should be implemented through the UN Environmental Program to give it global reach. By banning glyphosate,

Published by AAAS

we can avoid further irreversible ecosystem changes.

To make a ban on harmful herbicides possible, we must focus on natural and ecological weedkilling alternatives, such as root exudates (organic compounds secreted by plant roots), crop rotation, mulch, herbicidal soaps, fatty acids, and industrial vinegar (*11*). Most important, we must move to less intensive farming practices to reduce the massive global use of herbicides (*12*).

Wanxi Peng[1], Su Shiung Lam[2,1], Christian Sonne[3,1]*
[1]School of Forestry, Henan Agricultural University, Zhengzhou, China. [2]Universiti Malaysia Terengganu, Terengganu, Malaysia. [3]Aarhus University, Roskilde, Denmark.
*Corresponding author. Email: cs@bios.au.dk

REFERENCES AND NOTES

1. N. M. van Straalen, J. Legler, *Science* **360**, 958 (2018).
2. C. M. Benbrook, *Environ. Sci. Eur.* **28**, 3 (2016).
3. R. F. Service, *Science* **316**, 1116 (2007).
4. W. M. Farina *et al.*, *Insects* **10**, 354 (2019).
5. M. L. Eng *et al.*, *Science* **365**, 1177 (2019).
6. H. Siviter *et al.*, *Nature* **561**, 109 (2018).
7. R. Carson, *Silent Spring* (Houghton Mifflin, Boston, 1962).
8. PAN Europe, "Austria takes the lead to ban glyphosate, bringing human health and the environment to the forefront" (2019); www.pan-europe.info/press-releases/2019/12/austria-takes-lead-ban-glyphosate-bringing-human-health-and-environment.
9. C. Sonne, A. K. O. Alstrup, *Science* **363**, 938 (2019).
10. D. Butler, *Nature* **555**, 150 (2018).
11. P. Catherine, P. Josep, *Trends Plant Sci.* **25**, 14 (2020).
12. L. G. Horlings, T. K. Marsden, *Glob. Environ. Change* **21**, 441 (2011).

10.1126/science.aba5642

Fund plant conservation to solve biodiversity crisis

In their Letter "Solve the biodiversity crisis with funding" (20 September 2019, p. 1256), J. Malcom *et al.* called on the U.S. Congress to fund wildlife conservation programs to protect biodiversity. We agree that such funding is critically important, but we were disappointed that their discussion of biodiversity did not include plants. Even when unintentional, as this omission likely was, citing only animal examples can perpetuate the perception that plant conservation is less important and less worthy of funding (*1*). Plant conservation programs have been consistently underfunded, especially when compared to funding for animals. Although more than half of the species listed under the federal Endangered Species Act are plants, they receive less than 5% of the total funding for endangered species recovery (*2*, *3*).

This problem is exemplified by the Recovering America's Wildlife Act, introduced in July 2019 (*4*) and supported by fish and wildlife conservation groups (*5*, *6*). The bill would substantially improve funding for State Wildlife Action Plans (SWAPs), which are among the most effective species conservation programs in the United States. Plants can be listed in SWAPs, but because of antiquated authorizing language, the primary grants that fund SWAPs may only be used to conserve animal species of greatest conservation need, not plants (*7*). The proposed legislation does not update this language and would allow continued neglect of imperiled plants in SWAPs.

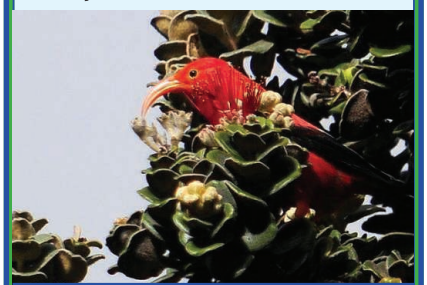
A recent global assessment found that at least 600 plant species are now extinct and that we are losing plant species at a rate 500 times higher than the extinction rate before human impacts (*8*). The conservation of plant species is essential to the successful conservation of fish, wildlife, pollinators, and other animals, as well as to human survival. It is critical that all efforts to improve funding for conservation explicitly include increased funding for plants. Without adequate conservation programs for plants, wildlife and biodiversity conservation efforts will inevitably be ineffective.

Emily Brin Roberson[1]*, Anne Frances[2,3,4], Kayri Havens[5,6], Joyce Maschinski[7], Abby Meyer[8], Lisa Ott[9,10]
[1]Native Plant Conservation Campaign, San Francisco, CA 94107, USA. [2]NatureServe, Arlington, VA 22202, USA. [3]Smithsonian-Mason School of Conservation, Front Royal, VA 22630, USA. [4]North American Plant Red List Authority, Arlington, VA 22202, USA. [5]Plant Conservation Alliance Nonfederal Cooperators Committee. [6]Chicago Botanic Garden, Glencoe, IL 60022, USA. [7]Center for Plant Conservation and San Diego Zoo Global Institute for Conservation Research, Escondido, CA 92027, USA. [8]Botanic Gardens Conservation International, U.S. at The Huntington Library, Art Museum, and Botanical Gardens, San Marino, CA 91108 USA. [9]The Garden Club of America, New York, NY 10022, USA. [10]North Shore Land Alliance, Oyster Bay, NY 11771 USA.
*Corresponding author.
Email: emilyr@plantsocieties.org

REFERENCES AND NOTES

1. M. Balding, K. J. H. Williams, *Conserv. Biol.* **30**, 1192 (2016).
2. K. Havens *et al.*, *Int. J. Plant Sci.* **175**, 3 (2014).
3. V. Negron-Ortiz, *Biol. Conserv.* **171**, 36 (2014).
4. 116th Congress, H.R.3742 - Recovering America's Wildlife Act of 2019 (2019); https://www.congress.gov/bill/116th-congress/house-bill/3742/text.
5. The Wildlife Society, Recovering America's Wildlife Act: A 21st century model of wildlife conservation funding (2019); https://wildlife.org/policy/recovering-americas-wildlife-act/.
6. American Fisheries Society, Recovering America's Wildlife Act: A significant, proactive conservation effort for imperiled species (2019); https://fisheries.org/policy-media/recovering-americas-wildlife-act/.
7. B. A. Stein, K. Gravuer, *Hidden in Plain Sight: The Role of Plants in State Wildlife Action Plans* (NatureServe Arlington, Virginia, 2008).
8. A. M. Humphreys *et al.*, *Nat. Ecol. Evol.* **3**, 1043 (2019).

10.1126/science.aba4360

AAAS Travels
Members and friends of AAAS are invited to join AAAS Travels on fascinating trips to all 7 continents!

Botanical HAWAII
April 1 - 11, 2020
Explore botanical gardens on the island of Oahu, the big island of Hawaii, and Kauai, including the geological marvels and native Hawaiian trees, plants, and birds at Hawaii Volcano National Park. From the brilliant red i'iwi bird to the Koa and Kukui trees, Hawaii is filled with natural wonders for you to discover!
$4,845 pp twin + air
Mystique of MOROCCO

April 3-15, 2020
Magical moments await you on this wonderful adventure to explore the mystique of the Kingdom of Morocco! Morocco possesses a wealth of opportunities for the traveler, spectacular landscapes and fascinating cultural experiences! Join us on this incredible adventure! $3,995 pp twin + air
For a detailed brochure, please call (800) 252-4910
All prices are per person twin share + air
Betchart Expeditions Inc.
17050 Montebello Rd
Cupertino, California 95014
Email: AAASInfo@betchartexpeditions.com
www.betchartexpeditions.com

Published by AAAS

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

BIOMECHANICS

Wing shapes take flight

Birds can dynamically alter the shape of their wings during flight, although how this is accomplished is poorly understood. Matloff *et al.* found that two mechanisms control the movement of the individual feathers. Whenever the skeleton moves, the feathers are redistributed passively through compliance of the elastic connective tissue at the feather base. To prevent the feathers from spreading too far apart, hook-shaped microstructures on adjacent feathers form a directional fastener that locks adjacent feathers. These features are found across a range of bird sizes; however, because the detachment of the hooks is noisy, they are notably absent in silent fliers, such as barn owls. —MSL

Science, this issue p. 293

An acrobatic midair altercation between male Baltimore orioles, *Icterus galbula*

MASS EXTINCTION

An impact with a dash of volcanism

Around the time of the end-Cretaceous mass extinction that wiped out dinosaurs, there was both a bolide impact and a large amount of volcanism. Hull *et al*. ran several temperature simulations based on different volcanic outgassing scenarios and compared them with temperature records across the extinction event. The best model fits to the data required most outgassing to occur before the impact. When combined with other lines of evidence, these models support an impact-driven extinction. However, volcanic gases may have played a role in shaping the rise of different species after the extinction event. —BG

Science, this issue p. 266

QUANTUM CRITICALITY

Spin-charge entanglement

Many physical properties follow characteristic scaling laws near quantum critical points, which are associated with phase transitions at absolute zero temperature. The material $YbRh_2Si_2$ has an antiferromagnetic quantum critical point, where spin-related properties are expected to follow such a scaling. Unexpectedly, Prochaska *et al*. found that charge fluctuations follow a critical scaling as well. The researchers fabricated high-quality thin films of $YbRh_2Si_2$ and used transmission spectroscopy to measure the optical conductivity of the film and infer the scaling. Their findings point to a highly entangled state of charge and spin, which may also be responsible for the strange-metal phase in this material. —JS

Science, this issue p. 285

CELL BIOLOGY

Ion fluxes resolve organellar volume

Animal cells continuously sample the surrounding medium, a feature accentuated in immune cells. Sampling is accomplished by trapping external medium into membrane-bound vesicles or vacuoles. These structures are promptly resolved, thus avoiding accumulation of endomembranes and volume expansion. In a variety of cultured cells, Freeman *et al*. found that this resolution entails conversion of spherical vacuoles into thin tubules, a process that involves marked changes in surface-to-volume ratio (see the Perspective by King and Smythe). Shrinkage of membrane-bound structures is driven by ion fluxes and subsequent osmotic transfer of water. Shriveled vacuoles attract curvature-sensing proteins that promote the extension of fine tubules. Ion channels thereby control membrane remodeling,

PHOTO: ALAN MURPHY/MINDEN PICTURES

Published by AAAS

enabling receptor recycling and proper routing of cellular cargo. —SMH

Science, this issue p.301; see also p. 246

MAMMALIAN EVOLUTION

Making a mammalian ear

Mammals have keen hearing owing to their complex inner ear. In our vertebrate ancestors, as in extant reptiles, the three bones that make up the inner ear were instead part of the jaw. Understanding the functional transition of these bones is challenging given their small and delicate nature. Mao *et al*. describe a new genus and species of stem therian mammal represented by six well-preserved specimens, seemingly caught as they slept huddled together (see the Perspective by Schultz). The unprecedented preservation reveals a clear transitional stage between the two very different functions of the bones. —SNV

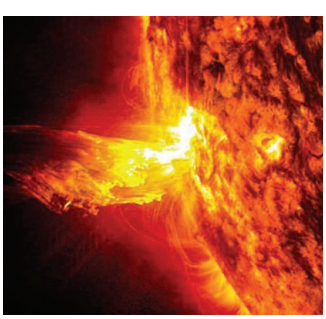
Science, this issue p. 305; see also p. 244

SOLAR PHYSICS

Magnetic energy release in a solar flare

Solar flares are bright flashes and associated eruptions of plasma from the Sun that are thought to be powered by violent rearrangement of the magnetic fields near sunspots. Fleishman *et al*. observed a bright solar flare with a microwave interferometer, allowing them to map the magnetic field in the solar corona and monitor how it changed during the flare. They found a large drop in the local field strength over 2 minutes, releasing enough magnetic energy to power the entire solar flare. Determining the origin of this energy will help to predict how strong future solar flares may be and their potential space weather impacts on Earth. —KTS

Science, this issue p. 278



Solar flare and accompanying ejected solar material

PHOTOS: (LEFT TO RIGHT) NASA/SDO; A. A. GONZALEZ *ET AL*., *NPJ SCI. LEARN.* **4**, 18 (2019)

ORGANOMETALLICS

Teaching bismuth to make and break bonds

One major reason why transition metals are good catalysts is that they can shuttle between oxidation states. This flexibility lets them slide in and out of chemical bonds; so, for instance, they can snip a bond between carbon and boron and then stitch a carbon-fluorine bond in its place. Planas *et al*. now report that bismuth can also orchestrate such bond-swapping events. They implement a fully catalytic cycle for fluorination of aryl boronates, in which bismuth hops between its +3 and +5 oxidation states. —JSY

Science, this issue p. 313

TRANSPLANTATION

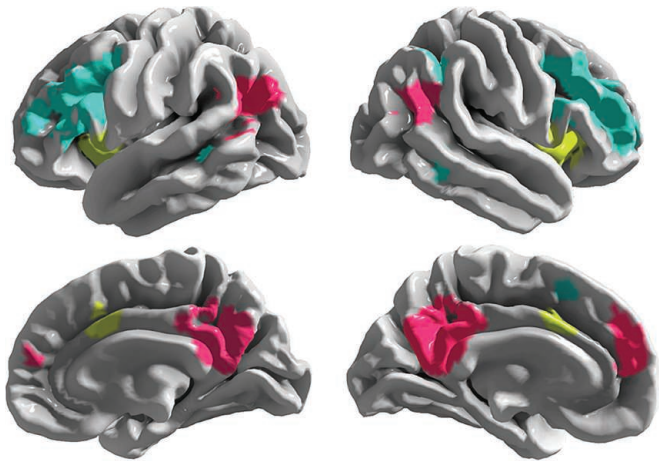
A better reception for islets

Pancreatic islet transplantation can stabilize blood glucose concentrations in individuals with type 1 diabetes, but transplant function decreases over time. Forbes *et al*. investigated the supportive effects of cotransplanting islets and human umbilical cord perivascular mesenchymal stromal cells (HUCPVCs). HUCPVCs inhibited T cells, expressed proregenerative and immune-regulatory markers in vitro, and increased islet vascularization in vivo. Cotransplanting human islets and HUCPVCs under the kidney capsule or via the hepatic portal vein improved control of blood glucose in diabetic immunodeficient and immunocompetent mouse models for up to 16 weeks. —CC

Sci. Transl. Med. **12**, eaan5907 (2020).

IN OTHER JOURNALS

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



EDUCATION

Sex, physics, and anxiety

Anxiety negatively affects academic performance, but little is known about the neural mechanisms underlying anxiety and STEM learning. To develop a more complete model of anxiety-related mechanisms and learning strategies, Gonzalez *et al*. evaluated anxiety and large-scale brain connectivity in 101 undergraduate physics students, collecting both self-reporting questionnaire and neuroimaging data. Results identified sex-specific relationships between STEM anxiety and brain connectivity, with male students exhibiting distinct internetwork connectivity for STEM and clinical anxiety and female students demonstrating no significant within-sex correlations. Using these data together with additional study results, the authors show that sex differences in brain networks are not fixed and that STEM anxiety is related to changes in both female and male students' brains during the physics-learning process. —MMc

NPJ Sci. Learn. **4**, 18 (2019).

Gender-related differences in neural mechanisms (within the colored brain regions) drive how anxiety is related to STEM learning.

FOREST ECOLOGY

The legacy of logging

Logging in tropical forests affects the future ecosystem functioning of the affected areas. Swinfield *et al*. investigated how logging affects the distribution of phosphorus, a key plant nutrient, over a wide area in Borneo. Spectroscopic imaging of forest canopies showed that the foliar concentrations of phosphorus were lower in logged than unlogged areas, indicating that soil phosphorus availability is decreased by logging. Because soil phosphorus is a key determinant of the tree species composition of tropical forests, the authors suggest that repeated logging of tropical forest on relatively infertile soils will lead to permanent, long-term changes in nutrient cycling and forest tree communities. —AMS

Glob. Chang. Biol. 10.1111/gcb.14903 (2019).

IMMUNOLOGY

Death suppression resolves inflammation

Impaired signaling of tyrosine-protein phosphatase

Published by AAAS

Individuals in a mixed species shoal of African cichlid fishes show distinct mate preferences that drive speciation.

SPECIATION

Picky mates drive speciation

Understanding how speciation occurs over small distances in the face of gene flow is a challenge. In these cases, local ecological adaptation can drive differentiation, but if gene flow across these regions is initially high, differentiation may not occur. Sibly *et al*. used a theoretical approach to explore the impact of a nonadaptive bias in mate choice on the potential for speciation in a system with local ecological differences, gene flow, and dispersal in both sexes. They suggest that mate preferences develop that can facilitate speciation. These can be in the form of imprinting, whereby an offspring develops a preference for mates that contain a trait like their parent, or matchmaking, whereby an individual prefers a trait that it itself has in its potential mate. These nonadaptive mating preferences, even in the face of dispersal, limit gene exchange among individuals adapted to local ecological conditions. Hence, selection for those conditions drives diversification. Such processes may be operating in habitats with concentrations of related species, such as cichlid fishes found in the lakes of the African Rift Valley. —SNV

Ecol. Evol. **9**, 13506 (2019).

non-receptor type 6 (PTPN6) has been linked to skin disorders. In neutrophilic dermatoses, inflammatory immune cells called neutrophils accumulate in the skin. PTPN6 normally plays a role in limiting inflammatory responses that are mediated through the interleukin-1 (IL-1α/β) cytokine receptor. Using mice with mutations in the *Ptpn6* gene, Speir *et al*. examined IL-1α/β release from neutrophils and asked how PTPN6 prevents inflammatory skin lesions. The researchers found that PTPN6 suppresses both apoptotic and necroptotic cell death in neutrophils, which in turn dampens IL-1–dependent inflammation. Controlling the nature and timing of neutrophil cell death in these diseases may therefore promote resolution of skin inflammation. —PNK

Nat. Immunol. **21**, 54 (2020).

GENE THERAPY

Correcting airways in cystic fibrosis

Cystic fibrosis is caused by inactivating mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene. Deletion of phenylalanine-508 is the most common mutation. Vaidyanathan *et al*. used CRISPR-associated protein 9 (Cas9)–mediated *CFTR* editing and adeno-associated virus delivery to correct the phenylalanine-508 deletion in upper-airway basal stem cells from 10 patients with cystic fibrosis. They achieved 30 to 50% gene correction and improved CFTR protein function. The corrected cells successfully engrafted within a clinically approved scaffold. This finding offers possibilities for clinical development of upper airway implants to treat respiratory failure, which is the biggest cause of mortality in cystic fibrosis patients. —GKA

Cell Stem Cell 10.1016/j.stem.2019.11.002 (2019).

ORGANOMETALLICS

Subtleties of cobalt oxidative addition

Chemical catalysis relies to a great extent on heavy precious transition metals that hop between oxidation states two electrons at a time. Recent efforts have focused on making greater use of the lighter, more abundant metals like iron, cobalt, and nickel. However, their tendency toward one-electron chemistry complicates the conventional mechanistic paradigms. Sandford *et al*. used a systematic combination of cyclic voltammetry and Hammett correlations of substituent effects to analyze the oxidative addition of a cobalt(I) complex to benzyl bromides. The results pointed to a two-step process involving a cobalt(II)-bromine intermediate and showcased the promising utility of the approach. —JSY

J. Am. Chem. Soc. **141**, 18877 (2019).

MATERIALS SYNTHESIS

Tracking replicate syntheses

Many synthetic routes are published each year for materials, but for how many of these materials is replicate synthesis reported? Agrawal *et al*. studied this question for metal-organic frameworks (MOFs), as these materials are often resynthesized to take advantage of the sorption properties or to make direct comparisons with new materials. They randomly selected 130 MOFs from a database of more than 4700 such materials and found that reported replicate syntheses were few and that the number diminished as a power law. Separately, they identified six MOFs that were resynthesized at rates far in excess of the power law prediction (some several hundred times). They also tracked the variation in their reported adsorption capacity as a function of years since the original publication. —PDS

Proc. Natl. Acad. Sci. U.S.A. 10.1073/pnas.1918484117 (2019).

PHOTO: RADEK BOROVKA/SHUTTERSTOCK

Published by AAAS

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

MICROBIOLOGY

A communicable component?

Noncommunicable diseases—such as metabolic disorders, cardiovascular disease, and cancer—are increasingly prevalent. Data suggest that these disorders are often associated with dysregulated gut microbiota (dysbiosis). In a Perspective, Finlay *et al*. propose that dysbiosis may be transmissible between individuals and that the etiology of noncommunicable diseases may actually have a communicable component. Although much more research is needed to ascertain whether microbes have a causative role in noncommunicable diseases, it is an important avenue of research that could have substantial implications for disease prevention and treatment. —GKA

Science, this issue p. 250

QUANTUM MATERIALS

An overview of an exotic type of liquid

Materials with interacting quantum spins that nevertheless do not order magnetically down to the lowest temperatures are candidates for a materials class called quantum spin liquids (QSLs). QSLs are characterized by long-range quantum entanglement and are tricky to study theoretically; an even more difficult task is to experimentally prove that a material is a QSL. Broholm *et al*. take a broad view of the state of the field and comment on the upcoming challenges. —JS

Science, this issue p. 263

T CELLS

A VISTA on naïve T cell fate

T cell quiescence and tolerance restrain the immune system from becoming overactive and attacking healthy tissue. Negative checkpoint regulators normally limit T cell responses to help safeguard against conditions such as autoimmunity. ElTanbouly *et al*. report that the checkpoint regulator VISTA (V-type immunoglobulin domain-containing suppressor of T cell activation) restricts early stages of T cell activation by shaping the inherent heterogeneity of the naïve $CD4^+$ T cell compartment to one that is more uniformly quiescent and silent (see the Perspective by Brown and Rudensky). Therapeutic targeting of VISTA using an agonistic antibody in mice curbed the development of graft-versus-host disease and promoted the death of naïve T cells abnormally activated by self-antigen. VISTA thus represents a distinctive immunoregulatory molecule that controls naïve T cell function by maintaining quiescence and peripheral tolerance. —PNK

Science, this issue p. 264; see also p. 247

ADVANCED IMAGING

Visualizing whole cells at many scales

Cells need to compartmentalize thousands of distinct proteins, but the nanoscale spatial relationship of many proteins to overall intracellular ultrastructure remains poorly understood. Correlated light and electron microscopy approaches can help. Hoffman *et al*. combined cryogenic super-resolution fluorescence microscopy and focused ion beam–milling scanning electron microscopy to visualize protein-ultrastructure relationships in three dimensions across whole cells. The fusion of the two imaging modalities enabled identification and three-dimensional segmentation of morphologically complex structures within the crowded intracellular environment. The researchers observed unexpected relationships within a variety of cell types, including a web-like protein adhesion network between juxtaposed cerebellar granule neurons. —SMH

Science, this issue p. 265

PALEONTOLOGY

A finer record of biodiversity

We have pressing, human-generated reasons to explore the influence of environmental change on biodiversity. Looking into the past can not only inform our understanding of this relationship but also help us to understand current change. Paleontological records depend on fossil availability and predictive modeling, however, and thus tend to give us a picture with large temporal jumps, millions of years wide. Such a scale makes it difficult to truly understand the action of environmental forces on ecological processes. Enabled by a supercomputer, Fan *et al*. used machine learning to analyze a large marine Paleozoic dataset, creating a record with time intervals of only ~26,000 years (see the Perspective by Wagner). This fine-scale resolution revealed new events and important details of previously described patterns. —SNV

Science, this issue p. 272; see also p. 249

OPTICS

Enhancing optical nonlinearity

Intense pulses of light interacting with a dielectric material can induce optical nonlinear behavior, whereby the frequency of the output light can be doubled or tripled or excited to even higher harmonics of the input light. Usually this interaction is weak and occurs over many thousands of wavelengths, typically requiring the combination of bulk volumes of material with a confining cavity. Using a mechanism of light confinement called bound states in the continuum, Koshelev *et al*. show that enhanced second-harmonic generation can be obtained in nanoscale subwavelength cylinders of a dielectric material. The results on these optical nanoantennas offer a platform for developing integrated nonlinear nanophotonic devices. —ISO

Science, this issue p. 288

ORGANIC CHEMISTRY

Sourcing nitrogen from nitromethane

Nitromethane is produced in bulk quantities for use as a solvent. Its applications as a reagent have focused mainly on the acidity of the methyl protons en route to modifying the carbon center. Liu *et al*. now report an alternative protocol that activates the nitrogen center to produce an aminating agent. An in situ reductive reaction with triflic anhydride, formic acid, and acetic acid yields an acetylated hydroxylamine, characterized by mass spectrometry. This nitrogen donor conveniently transforms a variety of ketones and aldehydes into amides. —JSY

Science, this issue p. 281

MARTIAN ATMOSPHERE

Water reaches Mars' upper atmosphere

Mars once hosted abundant water on its surface but subsequently lost most of it to space. Small amounts of water vapor are still present in the atmosphere, which can escape if they reach sufficiently high altitudes. Fedorova *et al*. used data from the ExoMars Trace Gas Orbiter spacecraft to determine the distribution of water in Mars' atmosphere and investigate how it varies over seasons. Water vapor is sometimes heavily saturated, and its distribution is affected by the planet's large dust storms.

Published by AAAS

Water can efficiently reach the upper atmosphere when Mars is in the warmest part of its orbit, and this behavior may have controlled the overall rate at which Mars lost its water. —KTS

Science, this issue p. 297

THERMAL CONDUCTIVITY

Thin graphite gets cool fast

In nonmetallic solids, heat is transported primarily through crystal vibrations called phonons. These phonons can have wavelike properties under certain conditions, which increases the thermal conductivity of the material. Machida *et al*. found that making graphite samples thin expands the hydrodynamic regime from cryogenic to room temperatures. The researchers measured an extremely high thermal conductivity in the very thin graphite samples, which may be important for a variety of electronics applications. —BG

Science, this issue p. 309

INFECTION

Why cholera is noninflammatory

Ingestion of food or water contaminated by the bacterium *Vibrio cholerae* results in cholera, which is a fatal disease if left untreated. *V. cholerae* releases toxins that translocate effector domains into infected cells, one of which induces cytoskeletal damage. Using human intestinal epithelial cells, Woida and Satchell found that proinflammatory signaling that would otherwise be activated by cytoskeleton damage was suppressed by other effector domains. These results may explain why cholera is a noninflammatory disease. —AMV

Sci. Signal. **13**, eaaw9447 (2020).

Published by AAAS

REVIEW SUMMARY

QUANTUM MATERIALS

Quantum spin liquids

C. Broholm, R. J. Cava, S. A. Kivelson, D. G. Nocera, M. R. Norman*, T. Senthil

BACKGROUND: Years ago, Lev Landau taught us how to think about distinct phases of matter through an order parameter that characterizes the symmetry-broken state relative to the symmetry-preserving state from which it emerges. More recently, however, it has been realized that not all phases of matter are captured by this paradigm. This was spectacularly demonstrated by the discovery of fractional quantum Hall states in the 1980s. Over the years, it has been elucidated that these states, along with their exotic excitations—quasiparticles carrying a rational fraction of the elementary charge of an electron—are the consequence of topological properties of ground state wave functions with a special type of long-range quantum entanglement. One might wonder whether analogous phenomena occur for spins. Whether these "quantum spin liquids" actually exist in nature has been the subject of much investigation.

ADVANCES: Since Philip Anderson contemplated the idea of quantum spin liquids in 1973, there has been a lot of research to establish what they are and how they can be characterized. Of particular note was the realization that an effective low-energy theory inevitably resembles the gauge theory treatments also invoked in high-energy physics. However, these gauge fields are "emergent" in the sense that they reflect important structure of the many-particle state. Specifically, they describe excitations that carry a fraction of the quantum of spin in terms of emergent quasiparticles with gauge charge and/or gauge flux, analogous to the electric charge and magnetic flux in electrodynamics. One consequence is that these quasiparticle excitations can have nontrivial statistical interactions when they are braided around each other. Although most studies have focused on gapped spin liquids, equally intriguing are gapless versions—for instance, ones where the quasiparticle ("spinon") spectrum is that of relativistic electrons described by the Dirac equation. Much work has been done to address specific models and connect them to experimental analogs. This has involved a combination of analytically solvable models, as well as the development of new numerical methods that provide approximate solutions given a microscopic (lattice scale) Hamiltonian.

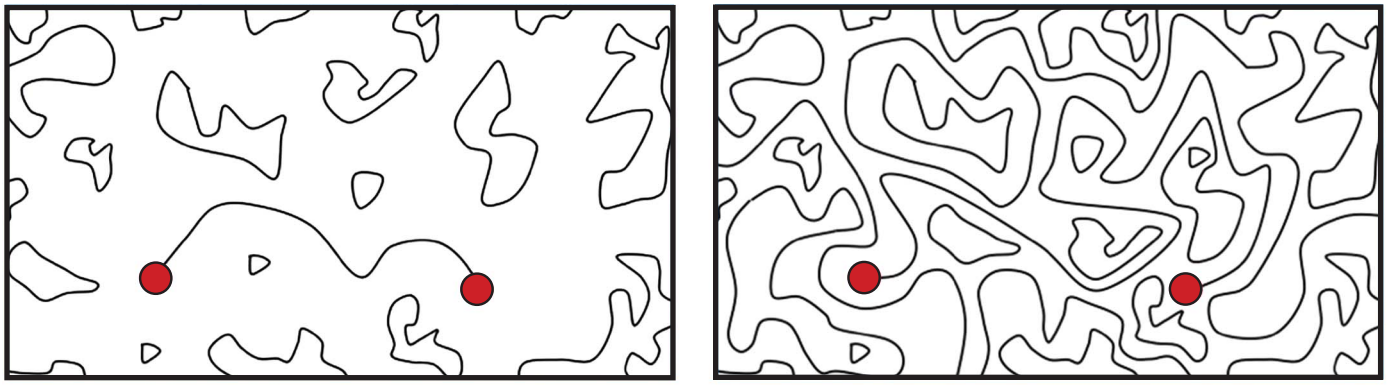
Perhaps most excitingly, there has been an increasingly promising effort to identify quantum spin liquids in nature. Much of the work has focused on materials where the magnetic ions reside on lattices that frustrate classical magnetic order. Examples include the triangular, kagome, hyperkagome, and pyrochlore lattices. Several candidate materials have been discovered, including organic salts, where molecular dimers realize spin-½ degrees of freedom on a distorted triangular lattice; herbertsmithite, where spin-½ copper ions form a kagome lattice; and α-$RuCl_3$, where j =1/2 ruthenium ions form a honeycomb lattice and that is thought to be proximate to the famous Kitaev model. All of these materials have properties reminiscent of spin liquids, though their documented fidelity as model systems is limited by disorder, subleading interactions, or lack of experimental information.

OUTLOOK: Given the infinite variety of potential materials and the many research groups now exploring this space, we are optimistic that a pristine materials realization of a quantum spin liquid will be discovered in the coming years. Perhaps even now a spin liquid exists in a long-forgotten drawer of a museum. Efforts to achieve ultrahigh-quality samples and new experiments designed to determine whether fractionalization and long-range entanglement occur in such materials will be key. In addition to tantalizing clues based on such techniques as thermal Hall conductivity, nuclear magnetic resonance, and inelastic neutron scattering, future methods may involve looking for spin currents to prove fractionalization, as has been done for charge degrees of freedom in the fractional quantum Hall case, or probing the range and character of quantum entanglement, as previously done in ultracold gases. Moreover, if quasiparticle excitations can be isolated and then manipulated, the prospect of a new form of topologically protected quantum computation also exists. Finally, chemically doped versions of spin liquids have been predicted to provide an unconventional route to superconductivity. The search for such phases will undoubtedly be an exciting undertaking. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aay0668

The list of author affiliations is available in the full article online.
*Corresponding author. Email: norman@anl.gov
Cite this article as C. Broholm *et al.*, *Science* **367**, eaay0668 (2020). DOI: 10.1126/science.aay0668



Emergent gauge theory as fluctuating loops. The loops are flux lines, with "particles" living at the ends of open lines. Left: The loops are dilute and small. The line connecting the particles costs a finite energy per unit length; the particles are confined. Right: The loops are numerous and include a fraction that are of macroscopic extent; the particles are free to move apart. This is the deconfined (spin liquid) phase.

REVIEW

QUANTUM MATERIALS

Quantum spin liquids

C. Broholm[1], R. J. Cava[2], S. A. Kivelson[3], D. G. Nocera[4], M. R. Norman[5]*, T. Senthil[6]

Spin liquids are quantum phases of matter with a variety of unusual features arising from their topological character, including "fractionalization"—elementary excitations that behave as fractions of an electron. Although there is not yet universally accepted experimental evidence that establishes that any single material has a spin liquid ground state, in the past few years a number of materials have been shown to exhibit distinctive properties that are expected of a quantum spin liquid. Here, we review theoretical and experimental progress in this area.

The history of spin liquids goes back to the early days of quantum mechanics. In 1928, Heisenberg achieved an understanding of ferromagnetism by considering a state in which all the spins point in a single direction (*1*). It is straightforward to see that a state of this sort is consistent with quantum mechanics (*2*). But problems emerged in considering antiferromagnets. Louis Néel's proposal that antiferromagnetism can be understood as a state in which the spins on alternating lattice sites point in alternating directions promoted great controversy at the time of its introduction; such a state cannot be the ground state (i.e., the lowest-energy state) of any reasonable quantum system (*3*, *4*). But it is now understood that the antiferromagnetic ground state is a prototypical example of the ubiquitous phenomenon of spontaneously broken global symmetry: The ground state is not spin-rotationally invariant and thus has a lower symmetry than the underlying Hamiltonian. This broken-symmetry point of view enables understanding of a number of universal properties of the antiferromagnetic state and their unity with similar phenomena in other ordered phases of matter. The same ideas when imported into particle physics underlie many of the successes of the standard model. For magnetic matter, it is now known that a variety of different kinds of spatially oscillating magnetic ordering patterns are possible, each corresponding to distinct broken symmetries. However, despite the successes of the broken-symmetry paradigm, the theoretical possibility of a "quantum spin liquid," for which there is no breaking of spin rotational symmetry, remained an intriguing possibility (*5*). In 1973 Philip Anderson proposed that the ground state of a simple quantum mechanical model—the spin-½ antiferromagnetic near-neighbor Heisenberg model (*6*) on a triangular lattice—might be a spin liquid. Specifically, he introduced the resonating valence bond (RVB) picture of a spin liquid wave function, based on the resonating single and double carbon-carbon bond picture developed by Linus Pauling and others to explain the electronic structure of benzene rings (*7*). Anderson's paper languished in relative obscurity until he resurrected the idea in the context of the high-temperature cuprate superconductors at the beginning of 1987 (*8*). It was realized soon afterward by Kivelson, Rokhsar, and Sethna (*9*) that the excitations of the spin liquid are topological in nature, and by Kalmeyer and Laughlin (*10*) that a version of the spin liquid could be constructed as a spin analog of the celebrated fractional quantum Hall state.

These developments in 1987 led to an explosion of interest in quantum spin liquids that continues to this day. In common with the fractional quantum Hall states, but distinct from conventional ordered states characterized by broken symmetry, the theory of the quantum spin liquid introduces new concepts, such as emergent gauge fields, into condensed-matter physics. It is not our intent here to cover the theory in great depth, as there exist several reviews (*11*–*15*). Rather, we wish to take a broader look at the field. In particular, what are the remaining big questions, both in theory and experiment?

What are quantum spin liquids?

To discuss them in the clearest context, let us focus on the idealized situation of quantum spins arranged in a periodic crystalline lattice, with interactions that are short-ranged in space. This setup describes correctly the essential physics of Mott (i.e., interaction driven) insulating materials. Mott insulating materials that do not magnetically order down to temperatures at which the spin dynamics is clearly quantum mechanical (i.e., much below the measured Curie-Weiss temperature) are attractive candidates in the search for spin liquids. However, this strategy is not sufficiently focused, as it includes nonmagnetic (quantum disordered) ground states that are not spin liquids (*16*, *17*). A more precise characterization comes from considering the structure of many-particle quantum entanglement in the ground state. A simple caricature of a magnetically ordered ground state wave function is achieved by specifying the spin on each site in the lattice. The ability to independently specify the quantum state of individual parts of a quantum many-particle system requires that the different parts have no essential quantum entanglement with each other. Thus, the prototypical ground state wave functions for conventional states of magnetic matter may be said to have short-range quantum entanglement between local degrees of freedom. By contrast, the quantum spin liquid refers to ground states in which the prototypical wave function has long-range quantum entanglement between local degrees of freedom (Fig. 1D). Under smooth deformations, such a wave function cannot be reduced to a product state wave function in real space (*18*). Such long-range quantum entanglement should be distinguished from the more familiar long-range order that characterizes broken-symmetry phases. Thus, the quantum spin liquid is a qualitatively new kind of ground state.

Just as there is no single type of magnetic order, there is no single type of quantum spin liquid. Loosely speaking, different types of quantum spin liquids correspond to different patterns of long-range entanglement. In addition, a useful (but coarse) classification distinguishes two classes of spin liquids on the basis of whether the excitation spectrum is separated from the ground state by an energy gap or not. Gapped spin liquids are simpler theoretically and are well characterized by the global topological structure of their ground state wave functions. Thus, they are said to have "topological order," a concept that also pertains to fractional quantum Hall systems. Such gapped spin liquids have well-defined emergent quasiparticles. These quasiparticles carry a topological signature that prevents them from being created in isolation (*9*, *12*). They can only be created in nontopological multiplets, which can then be pulled apart to yield multiple individual quasiparticles. A single isolated quasiparticle thus represents a nonlocal disturbance of the ground state. This nonlocality means that it can be detected far away by operations that involve moving other emergent quasiparticles. Thus, quasiparticle excitations have nonlocal "statistical" interactions (such as a charge moving around a magnetic flux). In two space dimensions, this

[1]Institute for Quantum Matter and Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA. [2]Department of Chemistry, Princeton University, Princeton, NJ 08544, USA. [3]Department of Physics, Stanford University, Stanford, CA 94305, USA. [4]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA. [5]Materials Science Division, Argonne National Laboratory, Argonne, IL 60439, USA. [6]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*Corresponding author. Email: norman@anl.gov

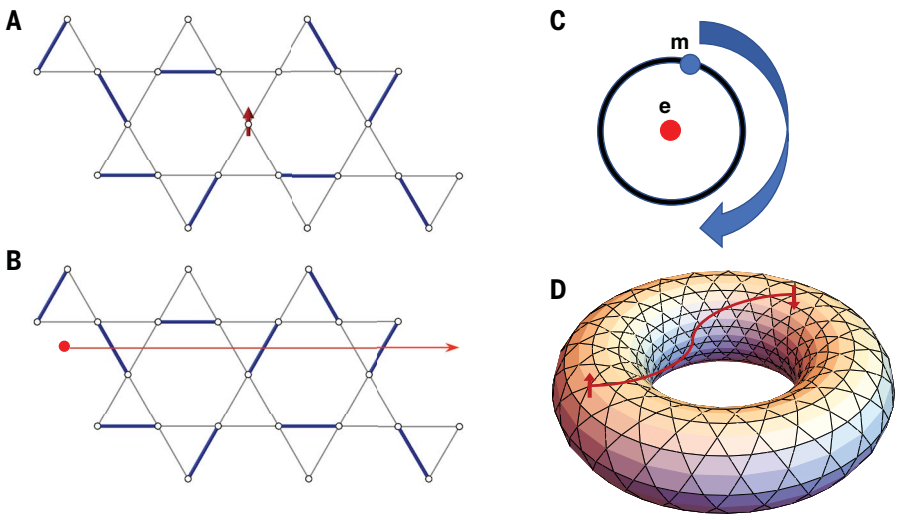


Fig. 1. Excitations of a spin liquid. Diagram of (**A**) a spinon excitation, (**B**) a vison excitation, and (**C**) braiding of anyons. Blue bonds represent spin singlets. The red arrow in (A) is a spinon, the red line with an arrow in (B) is a vison (where the phase of each singlet bond in the wave function intersected by this line changes sign), and e and m in (C) denote anyons. (**D**) Illustration of long-range entanglement of two spins, with the torus representing the ground state degeneracy typical for gapped spin liquids (the Z_2 spin liquid has a degeneracy of four on the torus associated with the topologically distinct horizontal and vertical loops that encircle the torus).

implies that the quasiparticles are "anyons" (*19*, *20*); that is, they pick up a nontrivial quantum-mechanical (Berry) phase when they circle around each other, as illustrated in Fig. 1C. This phase is associated with the "braiding" of the world lines traced by the quasiparticle trajectories.

There is a rich, formal theory of anyons in such topological ordered phases (*12*). In three space dimensions, in addition to emergent pointlike quasiparticles, there are also looplike excitations (analogous to flux lines in a superconductor) with a line tension. A quasiparticle encircling a loop excitation can also accrue a nontrivial phase. In either case (pointlike or looplike), the nonlocality associated with the quasiparticle excitation enables it to carry fractional quantum numbers associated with a global symmetry. A typical example of such a quasiparticle—known as a spinon—carries a spin of ½ and a charge of 0 (Fig. 1A). By contrast, local excitations in any insulating magnet must necessarily carry integer spin.

A second distinct class of spin liquids have a gapless excitation spectrum. In the simplest example of such a phase, the gapless spectrum admits a quasiparticle description. There also are gapless spin liquid phases where the quasiparticle description completely breaks down (*21*). In general, gapless spin liquids have power-law correlations of measurable quantities.

Given this variety of quantum spin liquid phases, what is the best theoretical framework in which we should think about them? Over the years, it has become clear that a powerful and convenient framework is provided by low-energy effective theories that involve emergent gauge fields (*22*–*25*), analogous to the vector potential in electrodynamics (*26*). Specifically, the low-energy effective theory of a quantum spin liquid is a deconfined gauge theory, that is, one in which spinons are free to propagate and thus not bound in pairs that would carry integer spin. (The particle physics analog would be a phase with free quarks.) The gauge theory description elegantly captures the nonlocal entanglement and its consequences.

To illustrate this, consider the case of a quantum spin liquid phase described by an emergent deconfined "Ising gauge field" (*27*–*30*), that is, a gauge field in which the magnetic flux can only take on two discrete values, 0 and 1. Formally, gauge theories are identified by their group structure—the Ising case is thus Z_2. Hence, this phase is known as a Z_2 quantum spin liquid. In two space dimensions, the excitations consist of a gapped excitation (the e "electric" particle) that carries Ising gauge charge and another gapped excitation (the m "magnetic" particle) that carries Ising gauge flux. These two excitations have a long-range statistical interaction: The wave function changes sign when an e particle is taken around an m particle (Fig. 1C). It is also possible to have a bound state of e and m (denoted ε). The e and m have bosonic statistics; however, their mutual braiding phase implies that ε has fermionic statistics. In systems with spin rotation symmetry, it can straightforwardly be shown that the e particle carries spin-½ (i.e., it is a spinon with bosonic statistics), whereas the m particle has a spin of 0; it is known as the "vison" (Fig. 1B). As their bound state, the ε particle also carries a spin of ½ and is known as the fermionic spinon (*31*, *32*).

There are multiple ways of thinking about how a phase with such an excitation structure might come about. A close and very useful analogy is with the excitations of the familiar Bardeen-Cooper-Schrieffer superconductor (*33*). The excitations of a superconductor include the Bogoliubov quasiparticle (resulting from the breaking up of a Cooper pair) and quantized vortices associated with $h/2e$ magnetic flux (here, h is Planck's constant and e is the electron charge). It is convenient to think about the quasiparticle in a basis where it is formally electrically neutral. In that instance, it has a braiding phase π with the $h/2e$ vortex. The Z_2 quantum spin liquid may be viewed as a phase-disordered version of a superconductor where long-range order is destroyed by quantum phase fluctuations. In this description, the fermionic spinon is identified as the cousin of the Bogoliubov quasiparticle (*26*, *34*, *35*), whereas the vison is identified as the cousin of the $h/2e$ vortex (*26*, *34*). The close relationship between the Z_2 spin liquid and the superconductor suggests that, if a spin liquid Mott insulator is found in a material, then doping it might naturally lead to superconductivity. Indeed, this is the original dream of the RVB theory as a mechanism for high-temperature superconductivity (*8*).

Other quantum spin liquid phases will have other emergent gauge groups, for example, the U(1) gauge field familiar from electromagnetism [U(1) being the group defined by rotations on a circle]; these are not obviously connected to superconductivity in any simple way. Given the importance of the gauge theory description, it is not surprising that many concepts in particle physics have been realized in the spin liquid context, including magnetic monopole-like excitations, which have been proposed in the context of the three-dimensional (3D) pyrochlore lattice (*36*). Furthermore, it is conceptually straightforward to combine features of a spin liquid with more conventional phases, giving rise to additional new quantum phases of matter with combined topological order and broken symmetries (*37*, *38*), or even new metallic phases with a Fermi surface whose enclosed volume violates Luttinger's theorem (that is, it is not proportional to the electronic density) (*39*).

Do quantum spin liquids exist in theory?

This question was settled in a variety of different ways in the late 1980s and 1990s, when the first stable effective field theory descriptions of both the Z_2 quantum spin liquid (*26*–*29*) and a different time-reversal broken version (known as a chiral spin liquid) (*40*) were developed and

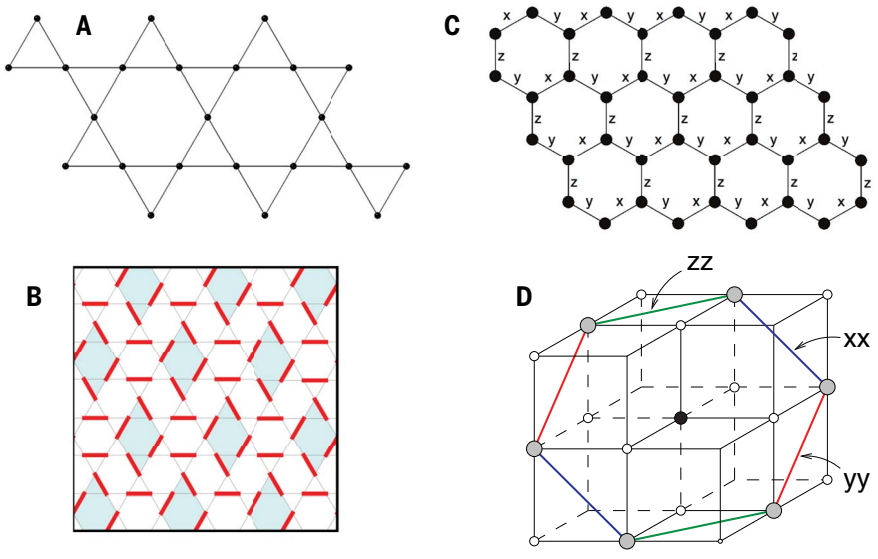


Fig. 2. Geometrically frustrated models. (**A**) Kagome lattice, (**B**) diamond valence bond solid on a kagome lattice (*153*), (**C**) Kitaev model on a honeycomb lattice, and (**D**) bond-dependent Kitaev interaction in a sixfold coordinated transition metal oxide (*67*). In (B), red bonds are singlets, with blue shading emphasizing the diamonds. In (C) and (D), x (xx), y (yy), and z (zz) denote the component of the spins involved in that bond.

their physical properties elucidated. Specific models that realize the Z_2 spin liquid were constructed in an SU(N) generalization of the SU(2) Heisenberg magnet on square lattices with short-range interactions involving more than just nearest neighbors (*28*) (so as to frustrate classical order) and on frustrated nonbipartite lattices [e.g., the triangular and kagome lattices (*41*)]. A Z_2 topological ordered state was also shown to be present in the quantum dimer model (*42*) on the triangular lattice (*43*). Additionally, Kitaev described a simple exactly solvable model (the toric code) for a Z_2 spin liquid (*44*). Building on these developments, many concrete models were constructed and reliably shown to have spin liquid phases with a variety of emergent gauge structures, in both two (*45*–*47*) and three dimensions (*46*, *48*). Though the matter of principle question has been answered in the affirmative, the question of which of these phases, if any, occur in realistic models of materials remained largely open and is still not satisfactorily settled.

Anderson's idea in 1973 that the ground state of the near-neighbor Heisenberg model was a spin liquid is not realized for the simplest form of the triangular lattice antiferromagnet, even for spin-½ systems where quantum effects are maximized, as was shown by Huse and Elser (*49*) among others. Modifications of the ideal model—for instance, the inclusion of ring exchange (*50*), further neighbor coupling (*51*), or spin anisotropy (*52*)—can, however, lead to spin liquid states (as we allude to below when talking about real materials such as the 2D organic ET and dmit salts). This led to the study of other lattices where antiferromagnetic interactions are more frustrated (i.e., act to suppress long-range magnetic order). The classic example in 2D is the lattice of corner-sharing triangles known as the kagome lattice (Fig. 2A). In the case of a near-neighbor classical Heisenberg model on a kagome lattice, continuous rotations of spins on certain clusters are possible at no energy cost (*53*–*55*), implying a large manifold of soft fluctuation modes that act to suppress order. This is particularly evident in exact diagonalization studies (*56*), which show a spectrum of states qualitatively different from the triangular lattice case, with a dense set of both singlet and triplet excitations extending to low energies. Such studies have been unable to definitively address whether the excitation spectrum for both singlets and triplets is gapped or not because of limitations of modern supercomputers [the largest lattice studied so far has been 48 sites (*57*)]. Researchers have addressed larger lattices by using approximate techniques based on quantum information–like methods, such as the density matrix renormalization group (DMRG) and various generalizations, including projected entangled pair states (PEPS) and the multiscale entanglement renormalization ansatz (MERA). The basic conclusion of such studies of the kagome lattice is that there are a number of states that have almost equal energies (*13*), including gapped Z_2 spin liquids, gapless spin liquids [so-called U(1) spin liquids where the spinons have a Dirac-like dispersion], and long-period valence bond solids. The spin liquid ground state implied by DMRG studies (*58*) appears to be a "melted" version of a 12-site valence bond solid that has a diamondlike structure, as shown in Fig. 2B, although some studies point to a U(1) gapless spin liquid instead (*59*). Exact diagonalization studies suggest that the ground state might break inversion symmetry or even be chiral in nature (*56*). Moreover, because the kagome lattice lacks a point of inversion symmetry between neighboring sites, this allows for Dzyaloshinskii-Moriya (DM) interactions that can qualitatively change the ground state relative to that of the Heisenberg model. Indeed, there are indications from simulations that the addition of DM interactions favors magnetic order (*60*–*62*).

In 2006, another exactly solvable model was reported by Kitaev (*63*). Based on a honeycomb lattice, the Hamiltonian is a less symmetric version of the Heisenberg model (*6*), where exchange on the "*x*" bonds of the honeycomb involves only S_xS_x, on the "*y*" bonds only S_yS_y, and on the "*z*" bonds only S_zS_z (Fig. 2C). Its ground state is a Z_2 spin liquid with a gapless spectrum of fermionic ε particles (known as Majoranas). Making the model anisotropic between the x, y, and z bonds preserves the exact solubility but gaps out the ε particle. Notably, the exact solution yields not just the ground state but the full spectrum of excitations. The manifold of states can be factored into flux sectors, with the flux referring to the product of the sign of the singlets around a hexagonal loop in the honeycomb (for the ground state, +1 for all hexagons). Flux excitations are precisely the visons mentioned above and are localized with a small energy gap. But the "unbound" Majorana is free to propagate and forms a dispersion that can be either gapped or gapless, depending on the ratio of the various J (J_x, J_y, J_z). The interaction of these low-energy visons with the Majoranas leads to a rather featureless spin excitation spectrum, as could be measured by neutrons (*64*). One consequence of this model is emergent fermionic statistics in the continuum of spin excitations as would be measured by Raman scattering (*65*). Even more noteworthy is the prediction of Majorana edge currents in a magnetic field, which would lead to quantization of the thermal Hall effect with a value half that expected for fermionic edge modes (*66*). Despite the seemingly contrived form of this model, it was pointed out by Jackeli and Khaliullin in 2009 (*67*) that the model might be physically realized in certain honeycomb (and "hyperhoneycomb") iridates and related materials such as α-$RuCl_3$ (Fig. 2D), which has led to an explosion of interest in both this model and those materials. This brings us to our next question.

Do quantum spin liquids really exist in nature?

Although a spin-½ antiferromagnetic chain is a 1D analog of a quantum spin liquid [and its

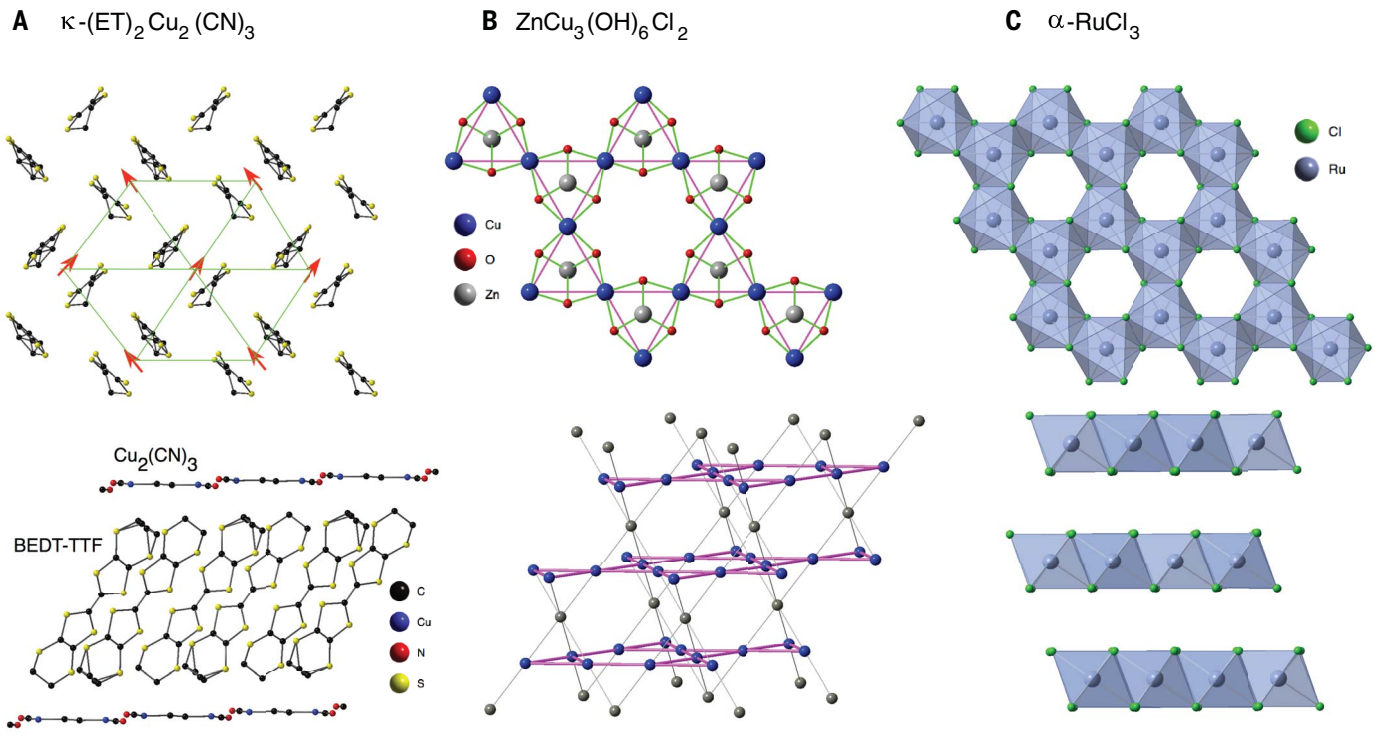


Fig. 3. Candidate spin liquid materials. Crystal structures of (**A**) κ-$(ET)_2Cu_2(CN)_3$, (**B**) herbertsmithite, and (**C**) α-$RuCl_3$. In (A), the ET dimers (top) form a triangular lattice (with the S = ½ spin degree of freedom per dimer represented by red arrows). These ET molecules are sandwiched by $Cu_2(CN)_3$ planes (bottom). In (B), Cu forms kagome layers (top) that are interconnected (bottom) by Zn (O is shown in the top only, and H and Cl have been suppressed). In (C), Ru octahedra (top) form honeycomb layers that are weakly coupled (bottom) with Cl.

spinons have been observed in experiments (*68*)], it is qualitatively different (for instance, there is no braiding in 1D). Beyond one dimension, a number of interesting candidate materials have emerged that might host quantum spin liquids, but the evidence is circumstantial. The focus has been on materials with spins on lattices that frustrate conventional Néel order. Spin-½ systems are of particular interest because they are the least classical, but the possibility of long-range entanglement for higher spin states should not be overlooked. Fluctuations are enhanced in 2D and for low coordination numbers, but even in 3D, there are pyrochlore and hyperkagome lattice systems that fail to develop magnetic order owing to geometrical frustration. Our theoretical understanding further suggests that "weak" Mott insulators that are close to the metal-insulator transition are fertile grounds for quantum spin liquid phases, consistent with the recent discovery of frustrated magnetism near the Mott transition in $(V_{1-x}Cr_x)_2O_3$ (*69*). Three of the most actively discussed classes of materials at the present time are shown in Fig. 3, and all involve lattices where either the spin, s, or the total angular momentum, j, has a value of ½. They are (i) 2D organic salts such as κ-$(ET_2)Cu_2(CN)_3$ and $EtMe_3Sb[Pd(dmit)_2]_2$, where structural dimers possessing a single spin-½ degree of freedom form a triangular lattice (*70*); (ii) herbertsmithite (and the closely related Zn-barlowite), where the Cu^{2+} ions form a kagome lattice (*71*); and (iii) α-$RuCl_3$, where the Ru^{3+} ions form a honeycomb lattice (*72*). The last two are deep in the Mott insulating phase, whereas the organic salts are weak Mott insulators close to the metal-insulator transition. We discuss each in turn, starting with the organics.

2D organic salts

Although most of these salts, where structural dimers form a (distorted) triangular lattice, have magnetic order at ambient pressure, there are a few that do not. Prominent examples are κ-$(BEDT\text{-}TTF)_2Cu_2(CN)_3$ (referred to here as κ-ET), κ-$(BEDT\text{-}TTF)_2Ag_2(CN)_3$, $EtMe_3Sb[Pd(dmit)_2]_2$ (referred to here as Pd-dmit), κ-$H_3(Cat\text{-}EDT\text{-}TTF)_2$, and κ-$(BEDT\text{-}TTF)_2Hg(SCN)_2Br$. Under pressure, κ-ET becomes superconducting, which was why it was first synthesized and studied (*73*). Nuclear magnetic resonance (NMR) studies show a lack of spin ordering down to temperatures well below the Curie-Weiss temperature inferred from high-temperature spin susceptibility measurements. At low temperatures, the spin susceptibility χ is a constant and the heat capacity $C = \gamma T$ has a linear temperature dependence (*74*). The Wilson ratio χ/γ is within 20% of the free Fermi gas value, which suggests that there are gapless spin-carrying excitations despite the lack of magnetic long-range order.

In Pd-dmit, despite its insulating nature, the thermal conductivity was reported to have a metallic form at low temperatures ($\kappa \propto T$) and is magnetic field dependent (*75*). If correct, this suggests that the gapless spin-carrying excitations are also mobile in this material. However, very recently this result has been reexamined in a number of dmit samples by two groups, and no such metallic thermal conductivity was found (*76*, *77*). Moreover, in κ-ET, there is at very low temperatures a dip in the thermal conductivity that, if taken at face value, might indicate a very small energy gap (*78*). This emphasizes the challenges associated with measurements of subtle features in complex materials with competing phases and the need for new results on spin liquid candidates to be thoroughly investigated. In κ-$(BEDT\text{-}TTF)_2Hg(SCN)_2Br$, heat capacity and Raman scattering indicate magnetic and electric dipole degrees of freedom that remain fluctuating to the lowest measured temperatures (*79*). Theoretically, the details of exactly which spin liquid is realized in these materials is not established. The experiments suggest that there may be a Fermi surface of emergent fermionic spinons (at least at very low temperatures). Establishing the presence of such a charge neutral Fermi surface in experiments

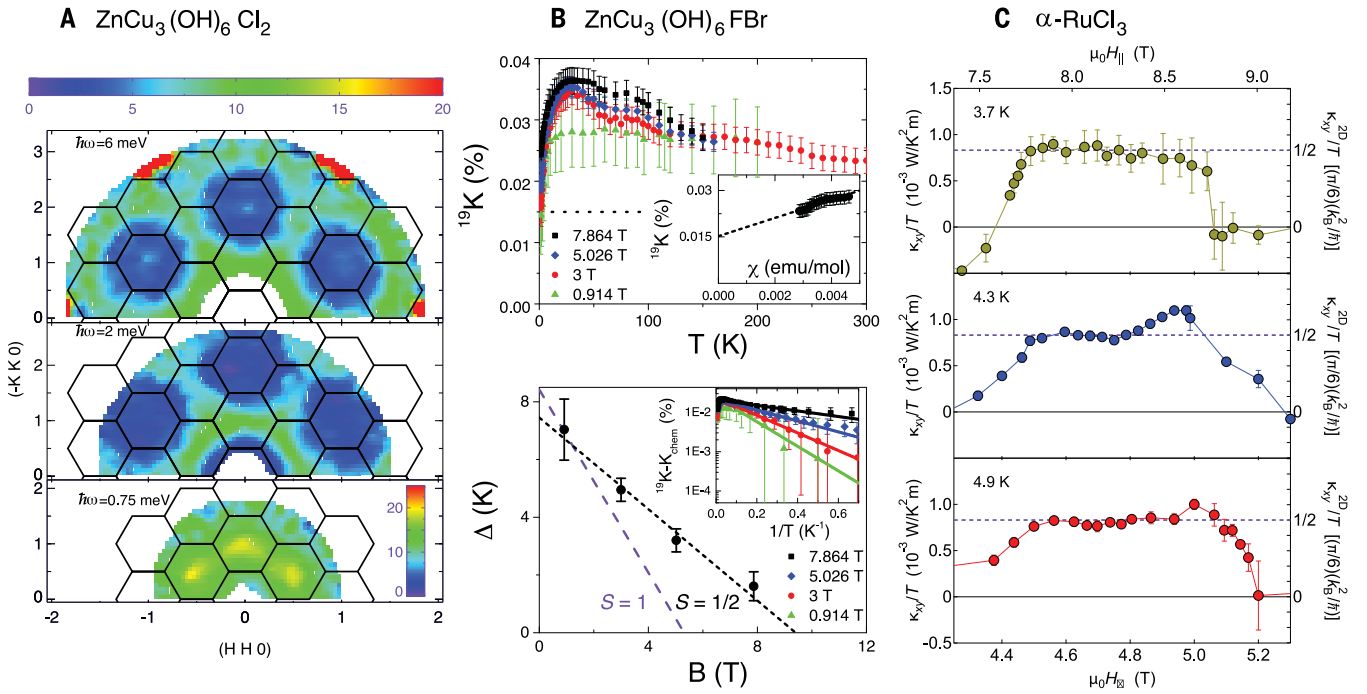


Fig. 4. Key data on spin liquid candidates. (**A**) Spin continuum of herbertsmithite from inelastic neutron scattering [S(**q**,ω) at 1.6 K in the *hk*0 plane: upper, 6 meV; middle, 2 meV; lower, 0.75 meV] (*90*). (**B**) Field dependence of the spin gap of Zn-barlowite from NMR [upper: ^{19}F Knight shift versus temperature for various magnetic fields; lower: magnetic field dependence of the spin gap, Δ, with dashed lines the expected behavior for S = ½ and S = 1 excitations] (*100*), and (**C**) quantized plateau in the thermal Hall effect of α-$RuCl_3$ [κ_{xy}/T versus magnetic field: upper, 3.7 K; middle 4.3 K; lower, 4.9 K] (*111*).

would be a great boost to our understanding. In that context, these materials (under pressure) exhibit quantum oscillations associated with their metallic Fermi surfaces. Such oscillations were found to be absent in the insulating spin liquid regime (*80*). This is contrary to the prediction that spin liquids with a spinon Fermi surface might host quantum oscillations due to weak coupling of the neutral spinons to charge fluctuations (*81*). Alternative interpretations of the data that invoke disorder to produce a heterogeneous gapless state (*82*–*84*) also deserve further experimental and theoretical exploration.

Herbertsmithite

Mineralogy has been used to inspire the search for spin liquids, making one wonder whether spin liquids are hiding in some long-forgotten desk drawer in a museum [as in the case of the first known naturally forming quasicrystal (*85*)]. The original studies (*86*) were on iron jarosites (and vanadium and chromium variants) where the magnetic ions form a perfect kagome lattice and where interesting behavior such as spin chirality has been observed (*87*). Unfortunately, these materials have long-range magnetic order, and the magnetic ions are not spin-½. However, owing to larger crystal fields and spin-orbit coupling, Ru^{3+} and Os^{3+} jarosites are candidates for a j = ½ kagome lattice, though synthesizing these minerals with 4d and 5d transition-metal ions presents an appreciable challenge.

It would be desirable to find a copper analog, given the large antiferromagnetic exchange J known to exist in copper oxides. However, the kagome ions in jarosites are 3+, and so cannot be formed with spin-½ Cu^{2+} except in diluted form. A related mineral class does contain Cu^{2+}: herbertsmithite, $ZnCu_3(OH)_6Cl_2$, a rare mineral first identified from a mine in Chile (*88*). The material was synthesized by using a hydrothermal method (*89*), and no evidence of long-range order was found. Since then, single crystals have been grown by using a refinement of the hydrothermal technique (*90*). This has allowed for single-crystal neutron scattering studies that have revealed a broad continuum of spin excitations (*91*) (Fig. 4A). Surprisingly, these excitations can be described by a dynamic magnetic correlation function of the "local" form $S(\mathbf{q},\omega) = f(\mathbf{q})\, g(\omega)$, reminiscent of the marginal Fermi liquid conjecture of Varma and colleagues (*92*). Such a form is not predicted by any known spin liquid model, though some resemblance to the data can be found in models where low-energy visons interact with the spinons, as mentioned earlier (*93*). This raises the important question of disorder. In particular, though it is claimed that the kagome spin excitations are gapped [as inferred from NMR (*94*) and neutron studies (*95*)], in reality, the entire low-energy spectrum is dominated by impurity spins (often referred to as "orphan" spins). These spins originate from the transition-metal sites between the kagome planes that are not completely inhabited by nonmagnetic Zn but also include magnetic Cu^{2+} (*96*). Similar issues exist when Zn is replaced by other 2+ ions such as Mg or Cd. Getting rid of these impurity spins is a major challenge, not only for herbertsmithite but for most spin liquid candidates where similar effects occur. This is important because some of the properties seen in herbertsmithite are reminiscent of random spin singlet states where there is a distribution of exchange energies J (*97*), and it has been claimed that the inelastic neutron scattering (INS) data can be understood in this way, as well (*98*). Such random singlet states are not quantum spin liquids (because their wave functions have a product form), even though they do exhibit quantum-critical–like scaling.

These issues have led to the study of related materials such as Zn-barlowite, which is similar to herbertsmithite except that the kagome layers are stacked differently (*99*, *100*). One advantage of Zn-barlowite is that the fluorine NMR line is simple, given its nuclear spin of ½. Analysis of these NMR data indicates a spin gap whose field dependence is consistent with a gas of spin-½ particles (i.e., spinons) (*101*) (Fig. 4B). This is further supported by INS studies, which indicate that the INS spin gap is twice that inferred by NMR (consistent

with the fact that INS measures spin-1 excitations, i.e., pairs of spinons) (*102*), though as in herbertsmithite (*93*, *94*), the low-energy properties of barlowite are dominated by defect spins. Most recently, attempts have been made to dope herbertsmithite to realize the long-sought "doped spin liquid" popularized by Anderson in 1987 (*8*). However, intercalating Li (*103*) or replacing Zn^{2+} by Ga^{3+} (*104*) leads to localized polarons [as confirmed by density functional calculations (*105*)], and thus no mobile carriers as in high-temperature superconducting cuprates (*106*). Even if polarons were not to occur, DMRG simulations predict Wigner crystallization of the doped carriers (*107*).

α-$RuCl_3$

The proposal by Jackeli and Khaliullin (*67*) that certain Mott-Hubbard systems with partially filled t_{2g}-levels and strong spin-orbit coupling might realize the Kitaev model led to an intense search. The first materials studied were those such as α-Na_2IrO_3 and α-Li_2IrO_3, where Ir^{4+} ions (with effective j = ½) form a honeycomb lattice. Although these materials exhibit long-range magnetic order, polarized resonant x-ray data show that bond-directional Kitaev interactions (Fig. 2D) indeed occur in this class of materials (*108*). This demonstrates why the recent discovery of a variant, $H_3LiIr_2O_6$, that does not exhibit long range order is important (*109*).

The realization that α-$RuCl_3$ has properties similar to those of the iridate honeycomb materials led to a huge growth in these studies. In α-$RuCl_3$, magnetic Ru is found on a honeycomb lattice between close-packed Cl planes (Fig. 3C). This material is relatively easy to grow in single-crystalline form and manipulate (as the layers are van der Waals coupled, they can be exfoliated). Also, the thermal neutron absorption cross section for Ru is a factor of 170 less than for Ir, so α-$RuCl_3$ is amenable to INS studies, which reveal a continuum of spin excitations (*110*). However, there has been some debate about which properties of this material are attributable to the Kitaev model, as opposed to more traditional physics (stemming from the non-negligible Heisenberg interaction). In particular, questions have been raised whether magnon-like excitations could explain some (or all) of the data (*111*), given that the material does order at low temperatures. Nevertheless, the spin continuum as detected in Raman data seems to obey fermionic statistics (*65*). Most notably, magnetic order is suppressed upon applying a magnetic field, implying that a spin liquid phase might exist in a range of magnetic fields. This led to a measurement of a thermal Hall signal that plateaued in a small range of temperature and magnetic field (*112*) (Fig. 4C). The value of this plateau is consistent with Majorana edge modes, being one-half of the value for fermionic edge modes (*66*). The observation of such a quantized plateau is peculiar, given that the thermal Hall angle is small (the longitudinal thermal conductivity is dominated by phonons), but this has been explained by two different theory efforts (*113*, *114*). As with most important experiments in this field, this result has yet to be reproduced by other groups. In addition, consistent with the organics and herbertsmithite, disorder should play an important role as well, particularly given the presence of stacking faults (*115*). Finally, based on the evidence that α-$RuCl_3$ exhibits spin liquid behavior, it is of great interest to study the physical properties of electron- and hole-doped variants (*116*, *117*).

A big question looms for the honeycomb-based spin liquid candidates: Is the Kitaev model actually relevant to these materials? The spin liquid in the exact solution may have only a tiny regime of stability beyond the solvable limit, on the basis of numerical calculations of the Kitaev model supplemented with Heisenberg exchange interactions (*118*). Furthermore, in the exact solution, the vison gap is very small (only a few percent of the Kitaev exchange) and so thermally, the spin liquid state only occurs at very low temperatures. Recent calculations suggest that a certain spin-anisotropic "symmetric exchange" enhances the stability of the exactly solved spin liquid (*119*). Alternatively, the possibility that any spin liquid that occurs in α-$RuCl_3$ or the iridates may not be smoothly connected to the Kitaev spin liquid must be kept in mind (*120*).

Other candidate materials

Space considerations preclude a detailed account of other spin liquid candidates. Of recent interest has been $YbMgGaO_4$, where the Yb ions form a triangular lattice, albeit with disorder on the nonmagnetic cation site. It is easy to grow and study, and the small energy scales associated with the 4f Yb ion make it more amenable to certain types of studies [extensive neutron scattering studies have been done (*121*)]. It, too, has been claimed to possibly have a "spinon" Fermi surface (*122*), but as with most spin liquid candidates, disorder plays an important role (*82*, *123*)—in this case, Mg and Ga interchanges that distort the Yb environment (*124*). Similar considerations apply to $Ba_3CuSb_2O_9$ (*125*), where Cu/Sb interchanges occur. Another candidate, $Ca_{10}Cr_7O_{28}$, can be described as a triangular lattice of six Cr^{5+}-based spin-½ clusters—each consisting of an antiferromagnetic and a ferromagnetic triangle interacting ferromagnetically with each other. Extensive experimental and numerical work on this bilayer kagome material has established its spin Hamiltonian and the lack of static spin ordering at temperatures as low as 0.3 K (*126*). For both $YbMaGaO_4$ (*127*) and $Ca_{10}Cr_7O_{28}$ (*128*), however, the absence of a linear term in the thermal conductivity argues against the existence of a spinon Fermi surface. Moreover, a lack of long-range magnetic order has been reported in the triangular-based materials $NiGa_2S_4$ (*129*), $Ba_8CoNb_6O_{24}$ (*130*), $NaYbO_2$ (*131*), and $Ba_4NbIr_3O_{12}$ (*132*), as well as in the honeycomb-based material $BaCo_2As_2O_8$ (*133*). Recently, a copper oxide, averievite [$Cu_5V_2O_{10}$(CsCl)], was identified in which the copper ions form a pyrochlore slab. First discovered in a volcano in Kamchatka, the material was synthesized and subsequently languished in an academic thesis, only to be "rediscovered" (thanks to Google Scholar) (*134*). Substitution by zinc likely replaces the intersite copper ions (as in herbertsmithite), isolating the copper kagome layers, and the resulting susceptibility and specific heat are reminiscent of herbertsmithite (*132*). Several materials are also known where magnetic ions form a "hyperkagome" lattice (obtained by taking the kagome layer and pulling it into the third dimension). Of particular note are $Na_4Ir_3O_8$ (*135*) and $PbCuTe_2O_6$ (*136*), but again both have quenched disorder (for the former material, caused by partial occupation of the Na sites) and distortions (for the latter material, there are many exchange parameters associated with its distorted hyperkagome lattice). As for other frustrated 3D lattices, extensive studies on rare-earth and transition-metal pyrochlores are beyond the scope of this article, and the reader is referred to a recent review (*137*). An exciting recent proposal (*138*), which has received some experimental support (*139*, *140*), is that the layered transition-metal dichalcogenide 1T-TaS_2 might be a quantum spin liquid.

The future

This review of quantum spin liquids may leave one to ask, "What else is out there?" Almost certainly, a lot. As for materials, many interesting ones known in mineralogical form have yet to be made in the lab and studied for their magnetic properties. As an example, quetzalcoatlite (named after an Aztec god) has copper ions on a perfect kagome lattice (*141*). But it, like many other minerals, is known only by its structure and nothing else. A systematic study of potentially frustrated magnetism in mineral collections might be a good start, followed by attempts to make cleaner synthetic versions of the most-promising minerals. A recurrent challenge with frustrated magnets is that chemical disorder acts at the "ultraviolet" scale, giving rise to orphan spins. Clearly, more attention (and resources) needs to be devoted to synthesis, both in developing promising new synthesis routes (high pressure, hydrothermal, molecular beam epitaxy, etc.) and finding ways to mitigate and control disorder. This is a difficult task, but it is useful to recall that it took decades

of materials research to develop elemental silicon clean enough for applications and modulation-doped AlGaAs heterostructures that display the fractional quantum Hall effect. Disorder—especially when carefully controlled—can also be illuminating. The tantalizing possibility of replacing Fe by Ru or Os in jarosites has been mentioned above. Similarly, one wonders what the osmium analog of α-$RuCl_3$ would be like (*142*). And, of course, obtaining a doped spin liquid that is metallic would be the holy grail for many (*8*, *9*, *22*, *143*). This could potentially be accomplished by ionic liquid gating to avoid chemical disorder.

Having addressed materials-based issues above, we turn to theory. Although great strides have been made in numerical techniques, we still do not know, for instance, what the ground state is of the near-neighbor Heisenberg model on a kagome lattice, and less about many other frustrated lattices, or for "real" Hamiltonians that contain multiple exchange parameters as well as anisotropic exchange and DM terms. Still less is known about dynamical and non-equilibrium properties. Although neutron scattering when combined with theoretical calculations of the magnetic structure factor $S(\mathbf{q},\omega)$ can provide circumstantial evidence for a spin liquid (*144*), methods to probe entanglement are needed to obtain model-independent evidence. As spin liquids are spin relatives of the fractional quantum Hall effect, it would make sense to apply methods known from spintronics to search for spin currents (*145*, *146*), the spin Hall effect, spin noise, and other spin-related effects that might expose the nature of the spinons (if they indeed exist). As for visons, a proposal for their study was made many years ago (*147*) that involves looking for trapped magnetic flux in a spin-liquid ring. This experiment was actually performed on a superconducting cuprate with a null result (*148*), but obviously doing this sort of experiment on spin liquid candidate materials would be in order. Similarly, impurities can be exploited not only to trap Majorana fermions but also to induce Friedel oscillations near defects (that could be detected by a scanning tunneling probe) that could reveal a spinon Fermi surface should it exist (*149*). And tunneling has been advocated as a possible probe of how electrons could potentially fractionalize when injected into a spin liquid (*150*). Ultimately, if topological excitations were identified in a material, then one could think about probing and extending their phase coherence time and braiding them in steps toward their utilization for "topological" quantum computation (*151*) (Fig. 1C). As for other potential applications, we can think of no better way to end than with Michael Faraday's supposed response to William Gladstone's dismissal of a scientific discovery: "What use is it?" he quipped. "Why, sir, there is every probability that you will soon be able to tax it."

REFERENCES AND NOTES

1. W. Heisenberg, On the theory of ferromagnetism. *Z. Phys.* **49**, 619–636 (1928). doi: 10.1007/BF01328601
2. In particular, it is an eigenstate of the spin operator, S_z, where z is the direction toward which the spins point.
3. P. W. Anderson, Some memories of developments in the theory of magnetism. *J. Appl. Phys.* **50**, 7281–7284 (1979). doi: 10.1063/1.326937
4. I. Dyzaloshinskii, High T_c superconductivity: Band electrons vs. neutral fermions. *Phys. Scr.* **T27**, 89–95 (1989). doi: 10.1088/0031-8949/1989/T27/014
5. I. Ya. Pomeranchuk, Thermal conductivity of the paramagnetic dielectrics at low temperatures. *J. Phys. (USSR)* **4**, 357–374 (1941).
6. This model is described by the Hamiltonian $J\,\Sigma_{i,j}\,\mathbf{S}(i)\cdot\mathbf{S}(j)$, with the exchange interaction $J > 0$ and the sum is over sites i and j that are near neighbors.
7. P. W. Anderson, Resonating valence bonds: A new kind of insulator. *Mater. Res. Bull.* **8**, 153–160 (1973). doi: 10.1016/0025-5408(73)90167-0
8. P. W. Anderson, The resonating valence bond state in La_2CuO_4 and superconductivity. *Science* **235**, 1196–1198 (1987). doi: 10.1126/science.235.4793.1196; pmid: 17818979
9. S. A. Kivelson, D. S. Rokhsar, J. P. Sethna, Topology of the resonating valence-bond state: Solitons and high-T_c superconductivity. *Phys. Rev. B Condens. Matter* **35**, 8865–8868 (1987). doi: 10.1103/PhysRevB.35.8865; pmid: 9941277
10. V. Kalmeyer, R. B. Laughlin, Equivalence of the resonating-valence-bond and fractional quantum Hall states. *Phys. Rev. Lett.* **59**, 2095–2098 (1987). doi: 10.1103/PhysRevLett.59.2095; pmid: 10035416
11. L. Balents, Spin liquids in frustrated magnets. *Nature* **464**, 199–208 (2010). doi: 10.1038/nature08917; pmid: 20220838
12. L. Savary, L. Balents, Quantum spin liquids: A review. *Rep. Prog. Phys.* **80**, 016502 (2017). doi: 10.1088/0034-4885/80/1/016502; pmid: 27823986
13. Y. Zhou, K. Kanoda, T.-K. Ng, Quantum spin liquid states. *Rev. Mod. Phys.* **89**, 025003 (2017). doi: 10.1103/RevModPhys.89.025003
14. J. Knolle, R. A. Moessner, A field guide to spin liquids. *Annu. Rev. Condens. Matter Phys.* **10**, 451–472 (2019). doi: 10.1146/annurev-conmatphys-031218-013401
15. X.-G. Wen, Zoo of quantum-topological phases of matter. *Rev. Mod. Phys.* **89**, 041004 (2017). doi: 10.1103/RevModPhys.89.041004
16. C. K. Majumdar, D. K. Ghosh, On next-near neighbor interaction in a linear chain. *J. Math. Phys.* **10**, 1388–1398 (1969). doi: 10.1063/1.1664978
17. I. Affleck, T. Kennedy, E. H. Lieb, H. Tasaki, Rigorous results on valence-bond ground states in antiferromagnets. *Phys. Rev. Lett.* **59**, 799–802 (1987). doi: 10.1103/PhysRevLett.59.799; pmid: 10035874
18. More technically, the ground state wave function cannot be completely disentangled into a product state with a finite depth circuit made up of local unitary transformations.
19. F. Wilczek, Quantum mechanics of fractional-spin particles. *Phys. Rev. Lett.* **49**, 957–959 (1982). doi: 10.1103/PhysRevLett.49.957
20. By an anyon, we mean any quasiparticle that has nontrivial statistical interactions either with itself or with other quasiparticles; it is possible for an anyon to have bosonic self-statistics, and yet nontrivial mutual statistics with some other quasiparticle.
21. M. Hermele *et al*., Stability of U(1) spin liquids in two dimensions. *Phys. Rev. B Condens. Matter Mater. Phys.* **70**, 214437 (2004). doi: 10.1103/PhysRevB.70.214437
22. P. A. Lee, N. Nagaosa, X.-G. Wen, Doping a Mott insulator: Physics of high-temperature superconductivity. *Rev. Mod. Phys.* **78**, 17–85 (2006). doi: 10.1103/RevModPhys.78.17
23. E. Fradkin, *Field Theories of Condensed Matter Systems* (Addison-Wesley, 1991).
24. X. G. Wen, *Quantum Field Theory of Many-Body Systems* (Oxford Univ. Press, 2004).
25. S. Sachdev, *Quantum Phase Transitions* (Cambridge Univ. Press, 1999).
26. The quantum theory of charged particles is expressed in terms of their coupling to the vector and scalar potentials, even though it is only the electric and magnetic fields, which are derivatives of these gauge fields, that are directly observable. Similar considerations apply to emergent gauge fields.
27. T. Senthil, M. P. A. Fisher, Z_2 gauge theory of electron fractionalization in strongly correlated systems. *Phys. Rev. B Condens. Matter Mater. Phys.* **62**, 7850–7881 (2000). doi: 10.1103/PhysRevB.62.7850
28. R. Moessner, S. L. Sondhi, E. Fradkin, Short-ranged resonating valence bond physics, quantum dimer models, and Ising gauge theories. *Phys. Rev. B Condens. Matter Mater. Phys.* **65**, 024505 (2001). doi: 10.1103/PhysRevB.65.024504
29. N. Read, S. Sachdev, Large-N expansion for frustrated quantum antiferromagnets. *Phys. Rev. Lett.* **66**, 1773–1776 (1991). doi: 10.1103/PhysRevLett.66.1773; pmid: 10043303
30. X. G. Wen, Mean-field theory of spin-liquid states with finite energy gap and topological orders. *Phys. Rev. B Condens. Matter* **44**, 2664–2672 (1991). doi: 10.1103/PhysRevB.44.2664; pmid: 9999836
31. N. Read, B. Chakraborty, Statistics of the excitations of the resonating-valence-bond state. *Phys. Rev. B Condens. Matter* **40**, 7133–7140 (1989). doi: 10.1103/PhysRevB.40.7133; pmid: 9991099
32. S. Kivelson, Statistics of holons in the quantum hard-core dimer gas. *Phys. Rev. B Condens. Matter* **39**, 259–264 (1989). doi: 10.1103/PhysRevB.39.259; pmid: 9947147
33. M. Tinkham, *Introduction to Superconductivity* (McGraw-Hill, 1966).
34. S. A. Kivelson, D. S. Rokhsar, Bogoliubov quasiparticles, spinons, and spin-charge decoupling in superconductors. *Phys. Rev. B Condens. Matter* **41**, 11693–11696 (1990). doi: 10.1103/PhysRevB.41.11693; pmid: 9993614
35. T. H. Hansson, V. Oganesyan, S. L. Sondhi, Superconductors are topologically ordered. *Ann. Phys.* **313**, 497–538 (2004). doi: 10.1016/j.aop.2004.05.006
36. C. Castelnovo, R. Moessner, S. L. Sondhi, Spin ice, fractionalization, and topological order. *Annu. Rev. Condens. Matter Phys.* **3**, 35–55 (2012). doi: 10.1146/annurev-conmatphys-020911-125058
37. L. Balents, M. P. A. Fisher, C. Nayak, Dual order parameter for the nodal liquid. *Phys. Rev. B Condens. Matter Mater. Phys.* **60**, 1654–1667 (1999). doi: 10.1103/PhysRevB.60.1654
38. T. Senthil, M. P. A. Fisher, Fractionalization, topological order, and cuprate superconductivity. *Phys. Rev. B Condens. Matter Mater. Phys.* **63**, 134521 (2001). doi: 10.1103/PhysRevB.63.134521
39. T. Senthil, S. Sachdev, M. Vojta, Fractionalized fermi liquids. *Phys. Rev. Lett.* **90**, 216403 (2003). doi: 10.1103/PhysRevLett.90.216403; pmid: 12786577
40. X. G. Wen, F. Wilczek, A. Zee, Chiral spin states and superconductivity. *Phys. Rev. B Condens. Matter* **39**, 11413–11423 (1989). doi: 10.1103/PhysRevB.39.11413; pmid: 9947970
41. S. Sachdev, Kagomé- and triangular-lattice Heisenberg antiferromagnets: Ordering from quantum fluctuations and quantum-disordered ground states with unconfined bosonic spinons. *Phys. Rev. B Condens. Matter* **45**, 12377–12396 (1992). doi: 10.1103/PhysRevB.45.12377; pmid: 10001275
42. By quantum dimers, we mean a state described by a superposition of singlet bonds as discussed in D. S. Rokhsar and S. A. Kivelson (*152*).
43. R. Moessner, S. L. Sondhi, Resonating valence bond phase in the triangular lattice quantum dimer model. *Phys. Rev. Lett.* **86**, 1881–1884 (2001). doi: 10.1103/PhysRevLett.86.1881; pmid: 11290272
44. A. Yu. Kitaev, Fault-tolerant quantum computation by anyons. *Ann. Phys.* **303**, 2–30 (2003). doi: 10.1016/S0003-4916(02)00018-0
45. L. Balents, M. P. A. Fisher, S. M. Girvin, Fractionalization in an easy-axis kagome antiferromagnet. *Phys. Rev. B Condens. Matter Mater. Phys.* **65**, 224412 (2002). doi: 10.1103/PhysRevB.65.224412
46. T. Senthil, O. Motrunich, Microscopic models for fractionalized phases in strongly correlated systems. *Phys. Rev. B Condens. Matter Mater. Phys.* **66**, 205104 (2002). doi: 10.1103/PhysRevB.66.205104
47. O. I. Motrunich, T. Senthil, Exotic order in simple models of bosonic systems. *Phys. Rev. Lett.* **89**, 277004 (2002). doi: 10.1103/PhysRevLett.89.277004; pmid: 12513235
48. M. Hermele, M. P. A. Fisher, L. Balents, Pyrochlore photons: The U(1) spin liquid in a S=1/2 three-dimensional frustrated magnet. *Phys. Rev. B Condens. Matter Mater. Phys.* **69**, 064404 (2004). doi: 10.1103/PhysRevB.69.064404
49. D. A. Huse, V. Elser, Simple variational wave functions for two-dimensional Heisenberg spin-1/2 antiferromagnets.

Phys. Rev. Lett. **60**, 2531–2534 (1988). doi: 10.1103/PhysRevLett.60.2531; pmid: 10038378

50. G. Misguich, C. Lhuillier, B. Bernu, C. Waldtmann, Spin-liquid phase of the multiple-spin exchange Hamiltonian on the triangular lattice. *Phys. Rev. B Condens. Matter Mater. Phys.* **60**, 1064–1074 (1999). doi: 10.1103/PhysRevB.60.1064
51. Z. Zhu, S. R. White, Spin liquid phase of the S = 1/2 J_1-J_2 Heisenberg model on the triangular lattice. *Phys. Rev. B Condens. Matter Mater. Phys.* **92**, 041105 (2015). doi: 10.1103/PhysRevB.92.041105
52. Z. Zhu, P. A. Maksimov, S. R. White, A. L. Chernyshev, Topography of spin liquids on a triangular lattice. *Phys. Rev. Lett.* **120**, 207203 (2018). doi: 10.1103/PhysRevLett.120.207203; pmid: 29864346
53. I. Ritchey, P. Chandra, P. Coleman, Spin folding in the two-dimensional Heisenberg kagomé antiferromagnet. *Phys. Rev. B Condens. Matter* **47**, 15342–15345 (1993). doi: 10.1103/PhysRevB.47.15342; pmid: 10005922
54. T. Yildirim, A. B. Harris, Magnetic structure and spin waves in the Kagome jarosite compound $KFe_3(SO_4)_2(OH)_6$. *Phys. Rev. B Condens. Matter Mater. Phys.* **73**, 214446 (2006). doi: 10.1103/PhysRevB.73.214446
55. J. T. Chalker, P. C. W. Holdsworth, E. F. Shender, Hidden order in a frustrated system: Properties of the Heisenberg Kagomé antiferromagnet. *Phys. Rev. Lett.* **68**, 855–858 (1992). doi: 10.1103/PhysRevLett.68.855; pmid: 10046010
56. P. Lecheminant, B. Bernu, C. Lhuillier, L. Pierre, P. Sindzingre, Order versus disorder in the quantum Heisenberg antiferromagnet on the kagome lattice using exact spectra analysis. *Phys. Rev. B Condens. Matter* **56**, 2521–2529 (1997). doi: 10.1103/PhysRevB.56.2521
57. A. M. Läuchli, J. Sudan, R. Moessner, S=1/2 kagome Heisenberg antiferromagnet revisited. *Phys. Rev. B* **100**, 155142 (2019). doi: 10.1103/PhysRevB.100.155142
58. S. Yan, D. A. Huse, S. R. White, Spin-liquid ground state of the S = 1/2 kagome Heisenberg antiferromagnet. *Science* **332**, 1173–1176 (2011). doi: 10.1126/science.1201080; pmid: 21527676
59. Y.-C. He, M. P. Zaletel, M. Oshikawa, F. Pollmann, Signatures of Dirac cones in a DMRG study of the kagome Heisenberg model. *Phys. Rev. X* **7**, 031020 (2017). doi: 10.1103/PhysRevX.7.031020
60. M. Elhajal, B. Canals, C. Lacroix, Symmetry breaking due to Dzyaloshinsky-Moriya interactions in the kagome lattice. *Phys. Rev. B Condens. Matter Mater. Phys.* **66**, 014422 (2002). doi: 10.1103/PhysRevB.66.014422
61. O. Cépas, C. M. Fong, P. W. Leung, C. Lhuillier, Quantum phase transition induced by Dzyaloshinskii-Moriya interactions in the kagome antiferromagnet. *Phys. Rev. B Condens. Matter Mater. Phys.* **78**, 140405 (2008). doi: 10.1103/PhysRevB.78.140405
62. C.-Y. Lee, B. Normand, Y.-J. Kao, Gapless spin liquid in the kagome Heisenberg antiferromagnet with Dzyaloshinskii-Moriya interactions. *Phys. Rev. B* **98**, 224414 (2018). doi: 10.1103/PhysRevB.98.224414
63. A. Kitaev, Anyons in an exactly solved model and beyond. *Ann. Phys.* **321**, 2–111 (2006). doi: 10.1016/j.aop.2005.10.005
64. J. Knolle, D. L. Kovrizhin, J. T. Chalker, R. Moessner, Dynamics of a two-dimensional quantum spin liquid: Signatures of emergent Majorana fermions and fluxes. *Phys. Rev. Lett.* **112**, 207203 (2014). doi: 10.1103/PhysRevLett.112.207203
65. J. Nasu, J. Knolle, D. L. Kovrizhin, Y. Motome, R. Moessner, Fermionic response from fractionalization in an insulating two-dimensional magnet. *Nat. Phys.* **12**, 912–915 (2016). doi: 10.1038/nphys3809
66. J. Nasu, J. Yoshitake, Y. Motome, Thermal transport in the Kitaev model. *Phys. Rev. Lett.* **119**, 127204 (2017). doi: 10.1103/PhysRevLett.119.127204; pmid: 29341648
67. G. Jackeli, G. Khaliullin, Mott insulators in the strong spin-orbit coupling limit: From Heisenberg to a quantum compass and Kitaev models. *Phys. Rev. Lett.* **102**, 017205 (2009). doi: 10.1103/PhysRevLett.102.017205; pmid: 19257237
68. B. Lake, D. A. Tennant, C. D. Frost, S. E. Nagler, Quantum criticality and universal scaling of a quantum antiferromagnet. *Nat. Mater.* **4**, 329–334 (2005). doi: 10.1038/nmat1327; pmid: 15778717
69. J. C. Leiner *et al.*, Frustrated magnetism in Mott insulating $(V_{1-x}Cr_x)_2O_3$. *Phys. Rev. X* **9**, 011035 (2019). doi: 10.1103/PhysRevX.9.011035
70. B. J. Powell, R. H. McKenzie, Quantum frustration in organic Mott insulators: From spin liquids to unconventional superconductors. *Rep. Prog. Phys.* **74**, 056501 (2011). doi: 10.1088/0034-4885/74/5/056501
71. M. R. Norman, Herbertsmithite and the search for the quantum spin liquid. *Rev. Mod. Phys.* **88**, 041002 (2016). doi: 10.1103/RevModPhys.88.041002
72. H. Takagi, T. Takayama, G. Jackeli, G. Khaliullin, S. E. Nagler, Concept and realization of Kitaev quantum spin liquids. *Nat. Rev. Phys.* **1**, 264–280 (2019). doi: 10.1038/s42254-019-0038-2
73. U. Geiser *et al.*, Superconductivity at 2.8 K and 1.5 kbar in κ-(BEDT-TTF)$_2$Cu$_2$(CN)$_3$: The first organic superconductor containing a polymeric copper cyanide anion. *Inorg. Chem.* **30**, 2586–2588 (1991). doi: 10.1021/ic00012a005
74. S. Yamashita *et al.*, Thermodynamic properties of a spin-1/2 spin-liquid state in a κ-type organic salt. *Nat. Phys.* **4**, 459–462 (2008). doi: 10.1038/nphys942
75. M. Yamashita *et al.*, Highly mobile gapless excitations in a two-dimensional candidate quantum spin liquid. *Science* **328**, 1246–1248 (2010). doi: 10.1126/science.1188200; pmid: 20522768
76. P. Bourgeois-Hope *et al.*, Thermal conductivity of the quantum spin liquid candidate $EtMe_3Sb[Pd(dmit)_2]_2$: No evidence of mobile gapless excitations. *Phys. Rev. X* **9**, 041051 (2019).
77. J. M. Ni *et al.*, Absence of magnetic thermal conductivity in the quantum spin liquid candidate $EtMe_3Sb[Pd(dmit)_2]_2$ – revisited. *Phys. Rev. Lett.* **123**, 247204 (2019).
78. M. Yamashita *et al.*, Thermal-transport measurements in a quantum spin-liquid state of the frustrated triangular magnet κ-(BEDT-TTF)$_2$Cu$_2$(CN)$_3$. *Nat. Phys.* **5**, 44–47 (2009). doi: 10.1038/nphys1134
79. N. Hassan *et al.*, Evidence for a quantum dipole liquid state in an organic quasi–two-dimensional material. *Science* **360**, 1101–1104 (2018). doi: 10.1126/science.aan6286; pmid: 29880684
80. D. Watanabe *et al.*, Novel Pauli-paramagnetic quantum phase in a Mott insulator. *Nat. Commun.* **3**, 1090 (2012). doi: 10.1038/ncomms2082; pmid: 23011144
81. O. I. Motrunich, Orbital magnetic field effects in spin liquid with spinon Fermi sea: Possible application to κ-(ET)$_2$Cu$_2$(CN)$_3$. *Phys. Rev. B Condens. Matter Mater. Phys.* **73**, 155115 (2006). doi: 10.1103/PhysRevB.73.155115
82. K. Watanabe, H. Kawamura, H. Nakano, T. Sakai, Quantum spin-liquid behavior in the spin-1/2 random Heisenberg antiferromagnet on the triangular lattice. *J. Phys. Soc. Jpn.* **83**, 034714 (2014). doi: 10.7566/JPSJ.83.034714
83. I. Kimchi, A. Nahum, T. Senthil, Valence bonds in random quantum magnets: Theory and application to $YbMgGaO_4$. *Phys. Rev. X* **8**, 031028 (2018). doi: 10.1103/PhysRevX.8.031028
84. K. Riedl, R. Valenti, S. Winter, Critical spin liquid versus valence bond glass in triangular lattice organic antiferromagnet. *Nat. Commun.* **10**, 2561 (2019). doi: 10.1038/s41467-019-10604-3; pmid: 31189897
85. L. Bindi, P. J. Steinhardt, N. Yao, P. J. Lu, Natural quasicrystals. *Science* **324**, 1306–1309 (2009). doi: 10.1126/science.1170827; pmid: 19498165
86. D. G. Nocera, B. M. Bartlett, D. Grohol, D. Papoutsakis, M. P. Shores, Spin frustration in 2D kagomé lattices: A problem for inorganic synthetic chemistry. *Chemistry* **10**, 3850–3859 (2004). doi: 10.1002/chem.200306074; pmid: 15316993
87. D. Grohol *et al.*, Spin chirality on a two-dimensional frustrated lattice. *Nat. Mater.* **4**, 323–328 (2005). doi: 10.1038/nmat1353; pmid: 15793572
88. R. S. W. Braithwaite, K. Mereiter, W. H. Paar, A. M. Clark, Herbertsmithite, $Cu_3Zn(OH)_6Cl_2$, a new species, and the definition of paratacamite. *Mineral. Mag.* **68**, 527–539 (2004). doi: 10.1180/0026461046830204
89. M. P. Shores, E. A. Nytko, B. M. Bartlett, D. G. Nocera, A structurally perfect S = (1/2) kagomé antiferromagnet. *J. Am. Chem. Soc.* **127**, 13462–13463 (2005). doi: 10.1021/ja053891p; pmid: 16190686
90. S. Chu, P. Muller, D. G. Nocera, Y. S. Lee, Hydrothermal growth of single crystals of the quantum magnets: Clinoatacamite, paratacamite, and herbertsmithite. *Appl. Phys. Lett.* **98**, 092508 (2011). doi: 10.1063/1.3562010
91. T.-H. Han *et al.*, Fractionalized excitations in the spin-liquid state of a kagome-lattice antiferromagnet. *Nature* **492**, 406–410 (2012). doi: 10.1038/nature11659; pmid: 23257883
92. C. M. Varma, P. B. Littlewood, S. Schmitt-Rink, E. Abrahams, A. E. Ruckenstein, Phenomenology of the normal state of Cu-O high-temperature superconductors. *Phys. Rev. Lett.* **63**, 1996–1999 (1989). doi: 10.1103/PhysRevLett.63.1996; pmid: 10040734
93. M. Punk, D. Chowdhury, S. Sachdev, Topological excitations and the dynamic structure factor of spin liquids on the kagome lattice. *Nat. Phys.* **10**, 289–293 (2014). doi: 10.1038/nphys2887
94. M. Fu, T. Imai, T.-H. Han, Y. S. Lee, Evidence for a gapped spin-liquid ground state in a kagome Heisenberg antiferromagnet. *Science* **350**, 655–658 (2015). doi: 10.1126/science.aab2120; pmid: 26542565
95. T.-H. Han *et al.*, Correlated impurities and intrinsic spin-liquid physics in the kagome material herbertsmithite. *Phys. Rev. B* **94**, 060409 (2016). doi: 10.1103/PhysRevB.94.060409
96. D. E. Freedman *et al.*, Site specific X-ray anomalous dispersion of the geometrically frustrated kagomé magnet, herbertsmithite, $ZnCu_{(3)}(OH)_{(6)}Cl_{(2)}$. *J. Am. Chem. Soc.* **132**, 16185–16190 (2010). doi: 10.1021/ja1070398; pmid: 20964423
97. S. Ma, C. Dasgupta, C. Hu, Random antiferromagnetic chain. *Phys. Rev. Lett.* **43**, 1434–1437 (1979). doi: 10.1103/PhysRevLett.43.1434
98. T. Shimokawa, K. Watanabe, H. Kawamura, Static and dynamical spin correlations of the S=1/2 random-bond antiferromagnetic Heisenberg model on the triangular and kagome lattices. *Phys. Rev. B Condens. Matter Mater. Phys.* **92**, 134407 (2015). doi: 10.1103/PhysRevB.92.134407
99. T.-H. Han, J. Singleton, J. A. Schlueter, Barlowite: A spin-1/2 antiferromagnet with a geometrically perfect kagome motif. *Phys. Rev. Lett.* **113**, 227203 (2014). doi: 10.1103/PhysRevLett.113.227203; pmid: 25494085
100. C. M. Pasco *et al.*, Single-crystal growth of $Cu_4(OH)_6BrF$ and universal behavior in quantum spin liquid candidates synthetic barlowite and herbertsmithite. *Phys. Rev. Mater.* **2**, 044406 (2018). doi: 10.1103/PhysRevMaterials.2.044406
101. Z. Feng *et al.*, Gapped spin-1/2 spinon excitations in a new kagome quantum spin liquid compound $Cu_3Zn(OH)_6FBr$. *Chin. Phys. Lett.* **34**, 077502 (2017). doi: 10.1088/0256-307X/34/7/077502
102. Y. Wei *et al.*, Evidence for a Z_2 topological ordered quantum spin liquid in a kagome-lattice antiferromagnet. arXiv:1710.02991 [cond-mat.str-el] (2017).
103. Z. A. Kelly, M. J. Gallagher, T. M. McQueen, Electron doping a kagome spin liquid. *Phys. Rev. X* **6**, 041007 (2016). doi: 10.1103/PhysRevX.6.041007
104. P. Puphal *et al.*, Tuning of a kagome magnet: Insulating ground state in Ga-substituted $Cu_4(OH)_6Cl_2$. *Phys. Status Solidi, B Basic Res.* **256**, 1800663 (2019). doi: 10.1002/pssb.201800663
105. Q. Liu *et al.*, Electron doping of proposed kagome quantum spin liquid produces localized states in the band gap. *Phys. Rev. Lett.* **121**, 186402 (2018). doi: 10.1103/PhysRevLett.121.186402; pmid: 30444389
106. We also note a word of caution with materials produced by such reduction approaches. In using strong reducing reagents on herbertsmithite, such as alkali metals, one needs to avoid the reduction of the protons to make hydrogen instead of reducing Cu^{2+}. Even small amounts of deprotonated hydroxides in herbertsmithite will result in structural distortions, therefore potentially breaking crucial symmetry that might be required for stabilizing quantum spin liquid behavior.
107. H.-C. Jiang, T. Devereaux, S. A. Kivelson, Holon Wigner crystal in a lightly doped kagome quantum spin liquid. *Phys. Rev. Lett.* **119**, 067002 (2017). doi: 10.1103/PhysRevLett.119.067002; pmid: 28949592
108. S. H. Chun *et al.* Direct evidence for dominant bond-directional interactions in a honeycomb lattice iridate Na_2IrO_3. *Nat. Phys.* **11**, 462–466 (2015). doi: 10.1038/nphys3322
109. K. Kitagawa *et al.*, A spin-orbital-entangled quantum liquid on a honeycomb lattice. *Nature* **554**, 341–345 (2018). doi: 10.1038/nature25482; pmid: 29446382
110. A. Banerjee *et al.*, Neutron scattering in the proximate quantum spin liquid α-$RuCl_3$. *Science* **356**, 1055–1059 (2017). doi: 10.1126/science.aah6015; pmid: 28596361
111. S. M. Winter *et al.*, Breakdown of magnons in a strongly spin-orbital coupled magnet. *Nat. Commun.* **8**, 1152 (2017). doi: 10.1038/s41467-017-01177-0; pmid: 29074965

112. Y. Kasahara *et al*., Majorana quantization and half-integer thermal quantum Hall effect in a Kitaev spin liquid. *Nature* **559**, 227–231 (2018). doi: 10.1038/s41586-018-0274-0; pmid: 29995863
113. Y. Vinkler-Aviv, A. Rosch, Approximately quantized thermal Hall effect of chiral liquids coupled to phonons. *Phys. Rev. X* **8**, 031032 (2018). doi: 10.1103/PhysRevX.8.031032
114. M. Ye, G. B. Halász, L. Savary, L. Balents, Quantization of the thermal Hall conductivity at small Hall angles. *Phys. Rev. Lett.* **121**, 147201 (2018). doi: 10.1103/PhysRevLett.121.147201; pmid: 30339425
115. H. B. Cao *et al*., Low-temperature crystal and magnetic structure of α-$RuCl_3$. *Phys. Rev. B* **93**, 134423 (2016). doi: 10.1103/PhysRevB.93.134423
116. A. Koitzsch *et al*., Nearest-neighbor Kitaev exchange blocked by charge order in electron-doped α-$RuCl_3$. *Phys. Rev. Mater.* **1**, 052001 (2017). doi: 10.1103/PhysRevMaterials.1.052001
117. It is worth noting that while doping α-$RuCl_3$ by chemical reduction or oxidation, one should always be careful to perform the reaction under an inert and non-nitrogen atmosphere because Ru complexes react with nitrogen.
118. J. G. Rau, E. K.-H. Lee, H.-Y. Kee, Spin-orbit physics giving rise to novel phases in correlated systems: Iridates and related materials. *Annu. Rev. Condens. Matter Phys.* **7**, 195–221 (2016). doi: 10.1146/annurev-conmatphys-031115-011319
119. M. Gohlke, G. Wachtel, Y. Yamaji, F. Pollmann, Y. B. Kim, Quantum spin liquid signatures in Kitaev-like frustrated magnets. *Phys. Rev. B* **97**, 075126 (2018). doi: 10.1103/PhysRevB.97.075126
120. Y.-D. Li, X. Yang, Y. Zhou, G. Chen, Non-Kitaev spin liquids in Kitaev materials. *Phys. Rev. B* **99**, 205119 (2019). doi: 10.1103/PhysRevB.99.205119
121. J. A. M. Paddison *et al*., Continuous excitations of the triangular-lattice quantum spin liquid $YbMgGaO_4$. *Nat. Phys.* **13**, 117–122 (2017). doi: 10.1038/nphys3971
122. Y. Shen *et al*., Evidence for a spinon Fermi surface in a triangular-lattice quantum-spin-liquid candidate. *Nature* **540**, 559–562 (2016). doi: 10.1038/nature20614; pmid: 27919078
123. Z. Zhu, P. A. Maksimov, S. R. White, A. L. Chernyshev, Disorder-induced mimicry of a spin liquid in $YbMgGaO_4$. *Phys. Rev. Lett.* **119**, 157201 (2017). doi: 10.1103/PhysRevLett.119.157201; pmid: 29077468
124. Y. Li *et al*., Crystalline electric-field randomness in the triangular lattice spin-liquid $YbMgGaO_4$. *Phys. Rev. Lett.* **118**, 107202 (2017). doi: 10.1103/PhysRevLett.118.107202; pmid: 28339219
125. S. Nakatsuji *et al*., Spin-orbital short-range order on a honeycomb-based lattice. *Science* **336**, 559–563 (2012). doi: 10.1126/science.1212154; pmid: 22556246
126. C. Balz *et al*., Physical realization of a quantum spin liquid based on a complex frustration mechanism. *Nat. Phys.* **12**, 942–949 (2016). doi: 10.1038/nphys3826
127. Y. Xu *et al*., Absence of magnetic thermal conductivity in the quantum spin-liquid candidate $YbMgGaO_4$. *Phys. Rev. Lett.* **117**, 267202 (2016). doi: 10.1103/PhysRevLett.117.267202; pmid: 28059548
128. J. M. Ni *et al*., Ultralow-temperature heat transport in the quantum spin liquid candidate $Ca_{10}Cr_7O_{28}$ with a bilayer kagome lattice. *Phys. Rev. B* **97**, 104413 (2018). doi: 10.1103/PhysRevB.97.104413
129. S. Nakatsuji *et al*., Spin disorder on a triangular lattice. *Science* **309**, 1697–1700 (2005). doi: 10.1126/science.1114727; pmid: 16151004
130. Y. Cui *et al*., Mermin-Wagner physics, (*H,T*) phase diagram, and candidate quantum spin-liquid phase in the spin-1/2 triangular-lattice antiferromagnet $Ba_8CoNb_6O_{24}$. *Phys. Rev. Mater.* **2**, 044403 (2018). doi: 10.1103/PhysRevMaterials.2.044403
131. M. M. Bordelon *et al*., Field-tunable quantum disordered ground state in the triangular-lattice antiferromagnet $NaYbO_2$. *Nat. Phys.* **15**, 1058–1064 (2019). doi: 10.1038/s41567-019-0594-5
132. L. T. Nguyen, R. J. Cava, Trimer-based spin liquid candidate $Ba_4NbIr_3O_{12}$. *Phys. Rev. Mater.* **3**, 014412 (2019). doi: 10.1103/PhysRevMaterials.3.014412
133. R. Zhong, T. Gao, N. P. Ong, R. J. Cava, Weak-field induced nonmagnetic state in a Co-based honeycomb. arXiv:1910.08577 [cond-mat.str-el] (2019).
134. A. S. Botana, H. Zheng, S. H. Lapidus, J. F. Mitchell, M. R. Norman, Averievite: A copper oxide kagome antiferromagnet. *Phys. Rev. B* **98**, 054421 (2018). doi: 10.1103/PhysRevB.98.054421
135. Y. Okamoto, M. Nohara, H. Aruga-Katori, H. Takagi, Spin-liquid state in the S=1/2 hyperkagome antiferromagnet $Na_4Ir_3O_8$. *Phys. Rev. Lett.* **99**, 137207 (2007). doi: 10.1103/PhysRevLett.99.137207; pmid: 17930633
136. S. Chillal *et al*., A quantum spin liquid based on a new three-dimensional lattice. arXiv:1712.07942 [cond-mat.str-el] (2017).
137. J. G. Rau, M. J. P. Gingras, Frustrated quantum rare-earth pyrochlores. *Annu. Rev. Condens. Matter Phys.* **10**, 357–386 (2019). doi: 10.1146/annurev-conmatphys-022317-110520
138. K. T. Law, P. A. Lee, 1T-TaS_2 as a quantum spin liquid. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6996–7000 (2017). doi: 10.1073/pnas.1706769114; pmid: 28634296
139. M. Klanjšek *et al*., A high-temperature quantum spin liquid with polaron spins. *Nat. Phys.* **13**, 1130–1134 (2017). doi: 10.1038/nphys4212
140. H. Murayama *et al*., Coexisting localized and itinerant gapless excitations in a quantum spin liquid candidate 1T-TaS_2. arXiv:1803.06100 [cond-mat.str-el] (2018).
141. P. C. Burns *et al*., Quetzalcoatlite: A new octahedral-tetrahedral structure from a 2 x 2 x 40 μm^3 crystal at the Advanced Photon Source-GSE-CARS facility. *Am. Mineral.* **85**, 604–607 (2000). doi: 10.2138/am-2000-0424
142. M. A. McGuire, Q. Zheng, J. Yan, B. C. Sales, Chemical disorder and spin-liquid-like magnetism in the van der Waals layered 5d transition metal halide $Os_{0.55}Cl_2$. *Phys. Rev. B* **99**, 214402 (2019). doi: 10.1103/PhysRevB.99.214402
143. I. I. Mazin *et al*., Theoretical prediction of a strongly correlated Dirac metal. *Nat. Commun.* **5**, 4261 (2014). doi: 10.1038/ncomms5261; pmid: 24980208
144. S. C. Morampudi, A. M. Turner, F. Pollmann, F. Wilczek, Statistics of fractionalized excitations through threshold spectroscopy. *Phys. Rev. Lett.* **118**, 227201 (2017). doi: 10.1103/PhysRevLett.118.227201; pmid: 28621969
145. C.-Z. Chen, Q. Sun, F. Wang, X. C. Xie, Detection of spinons via spin transport. *Phys. Rev. B Condens. Matter Mater. Phys.* **88**, 041405 (2013). doi: 10.1103/PhysRevB.88.041405
146. S. Chatterjee, S. Sachdev, Probing excitations in insulators via injection of spin currents. *Phys. Rev. B Condens. Matter Mater. Phys.* **92**, 165113 (2015). doi: 10.1103/PhysRevB.92.165113
147. T. Senthil, M. P. A. Fisher, Fractionalization in the cuprates: Detecting the topological order. *Phys. Rev. Lett.* **86**, 292–295 (2001). doi: 10.1103/PhysRevLett.86.292; pmid: 11177814
148. D. A. Bonn *et al*., A limit on spin-charge separation in high-T_c superconductors from the absence of a vortex-memory effect. *Nature* **414**, 887–889 (2001). doi: 10.1038/414887a; pmid: 11780056
149. D. F. Mross, T. Senthil, Charge Friedel oscillations in a Mott insulator. *Phys. Rev. B Condens. Matter Mater. Phys.* **84**, 041102 (2011). doi: 10.1103/PhysRevB.84.041102
150. M. Barkeshli, E. Berg, S. Kivelson, Coherent transmutation of electrons into fractionalized anyons. *Science* **346**, 722–725 (2014). doi: 10.1126/science.1253251; pmid: 25378617
151. C. Nayak, S. H. Simon, A. Stern, M. Freedman, S. Das Sarma, Non-abelian anyons and topological quantum computation. *Rev. Mod. Phys.* **80**, 1083–1159 (2008). doi: 10.1103/RevModPhys.80.1083
152. D. S. Rokhsar, S. A. Kivelson, Superconductivity and the quantum hard-core dimer gas. *Phys. Rev. Lett.* **61**, 2376–2379 (1988). doi: 10.1103/PhysRevLett.61.2376; pmid: 10039096
153. K. Hwang, Y. Huh, Y. B. Kim, Z_2 gauge theory for valence bond solids on the kagome lattice. *Phys. Rev. B Condens. Matter Mater. Phys.* **92**, 205131 (2015). doi: 10.1103/PhysRevB.92.205131

Acknowledgments: We thank S. Chernyshev, M. Hermanns, P. Lee, Y. Lee, and S. Sondhi for their comments. **Funding:** C.B. and R.J.C. were supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, through DE-SC-0019331. S.A.K. was supported by the National Science Foundation grant DMR-1608055. M.R.N. was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division. T.S. was supported by the National Science Foundation grant DMR-1608505, and partially through a Simons Investigator Award. **Competing interests:** The authors declare no competing interests.

10.1126/science.aay0668

RESEARCH ARTICLE SUMMARY

T CELLS

VISTA is a checkpoint regulator for naïve T cell quiescence and peripheral tolerance

Mohamed A. ElTanbouly*, Yanding Zhao*, Elizabeth Nowak, Jiannan Li, Evelien Schaafsma, Isabelle Le Mercier, Sabrina Ceeraz, J. Louise Lines, Changwei Peng, Catherine Carriere, Xin Huang, Maria Day, Brent Koehn, Sam W. Lee, Milagros Silva Morales, Kristin A. Hogquist, Stephen C. Jameson, Daniel Mueller, Jay Rothstein, Bruce R. Blazar, Chao Cheng†, Randolph J. Noelle†

INTRODUCTION: A central tenet of the clonal selection theory (CST), the current cornerstone of adaptive immunity, states that naïve T cells are selected on the basis of specificity to antigen and that the rest of the repertoire remains agnostic to antigen challenge. However, CST and more recent paradigms have not considered the need for negative regulatory signals to maintain quiescence of the unselected T cells as a pivotal aspect of maintaining clonal selection. Silencing of naïve T cells, or tolerance, also becomes critical upon encounter with self-antigen, whereby self-reactive clones cannot be allowed to develop effector function. These events are known to be interrelated because the loss of quiescence precipitates a loss in tolerance.

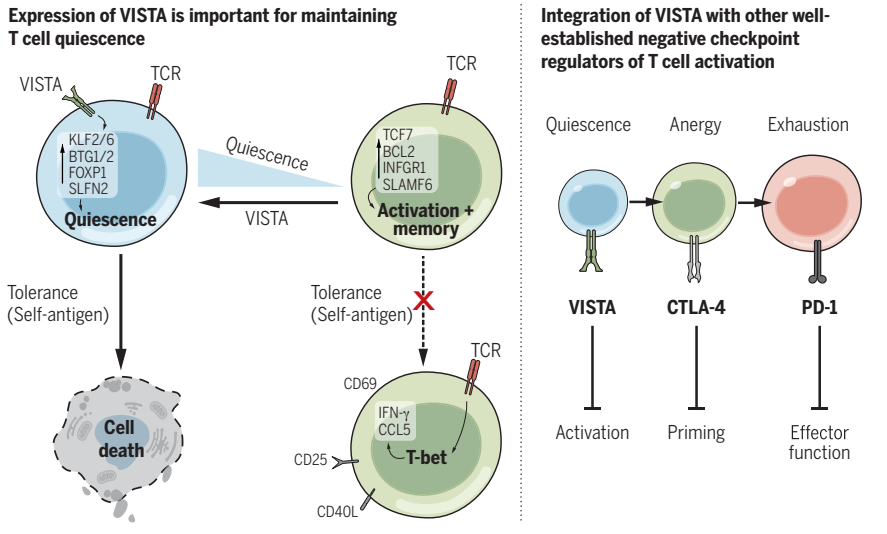
RATIONALE: V-type immunoglobulin domain-containing suppressor of T cell activation (VISTA) is an inhibitory receptor expressed on naïve T lymphocytes, and genetic deletion of VISTA culminates in T cell–mediated autoimmunity. Unlike other negative checkpoint regulators identified to date, VISTA is expressed on naïve T cells, which led us to investigate its function in maintaining naïve T cell quiescence and tolerance.

RESULTS: Using a combination of single-cell RNA sequencing and single-cell ATAC sequencing technologies coupled with in vivo analyses, we investigated the cell-state phenotype of naïve $CD4^+$ ($CD44^{lo}$ $CD62L^{hi}$) T cells in mice and found multiple T cell subpopulations exist within a surprisingly heterogeneous naïve T cell compartment. The vast majority of cells display a quiescent phenotype. However, there were subsets expressing less quiescent, memory-like features, a cluster with a high interferon-I (IFN-I) response signature, and a population with higher early T cell receptor (TCR) signaling genes. We show that genetic deletion of VISTA results in a major redistribution of the naïve T cell subsets with the notable reduction of the quiescent subset and the enhanced presence of a memory-like activated T cell subset. In the absence of intrinsic VISTA expression, naïve T cells were more epigenetically and transcriptionally primed toward stronger responses to TCR and cytokine stimulation, providing a rationale for the enhancement of $CD44^{hi}$ $CD4^+$ memory-like T cells observed in the absence of VISTA. As a consequence of reduced quiescence within the naïve T cell compartment, these naïve T cells were less susceptible to tolerance induction. Genetic deletion or antibody blockade of VISTA resulted in significant expansion and reduced tolerance of antigen-specific T cells. These effects were abolished under inflammatory conditions where VISTA expression on antigen-specific T cells was reduced. By contrast, agonistic engagement of VISTA and antigen exposure increased tolerance induction by enhancing the death and deletion of antigen-specific T cells. Tolerogenic death induced by anti-VISTA was demonstrated in a model of acute graft-versus-host disease (GVHD), whereby deletion of donor host reactive T cells resulted in a pronounced therapeutic benefit and long-term survival. The gene signature of $VISTA^{-/-}$ T cells overlapped significantly with T cells from patients with the autoimmune diseases systemic lupus erythematosus and rheumatoid arthritis, suggesting that VISTA may play a broad regulatory role in suppressing T cell self-reactivity.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aay0524

CONCLUSION: We introduce VISTA as the earliest checkpoint regulator of peripheral T cell tolerance identified to date and describe VISTA as the first of a new class of immunoregulatory molecules whose function is to enforce quiescence in naïve T lymphocytes. In addition, comprehensive epigenetic and transcriptional analyses defined fluorescence-activated sorting-purified $CD44^{lo}$ $CD62L^{hi}$ naïve $CD4^+$ T cells as heterogeneous. Our current in vivo findings also suggest that regulation of quiescence is inextricably linked to the capacity to maintain T cell self-tolerance. Although a critical player in naïve T cell homeostasis, the restraint enforced by VISTA on naïve T cell responses is lost when antigen stimulation is met under inflammatory conditions. ■



Loss of VISTA reduces the quiescence phenotype of the naïve T cells at the transcriptional and epigenetic levels, causing enhancement in TCR and cytokine pathways. (Left) This loss undermines tolerance induction to self-antigen and imparts a more inflammatory phenotype upon activation. KLF2/6, BTG1/2, FOXP1, and SLFN2 are regulators of quiescence. TCF7, BCL2, INFGR1, and SLAMF6 direct memory T cell fate and activity. IFN-γ and CCL5 are effector cytokines. CD69, CD25, and CD40L are activation and costimulation receptors. (Right) Integration of VISTA with other well-established negative checkpoint regulators of T cell activation. VISTA is constitutively expressed on naïve T cells and plays a critical role in enforcing quiescence. CTLA-4 is expressed on the cell surface briefly after T cell activation and inhibits T cell activation at the priming stage by limiting costimulation. PD-1 is expressed later during priming and inhibits T cells at the effector stage.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: rjn@dartmouth.edu (R.J.N.); chao.cheng@bcm.edu (C.C.)
Cite this article as M. A. ElTanbouly *et al.*, *Science* **367**, eaay0524 (2020). DOI: 10.1126/science.aay0524

RESEARCH ARTICLE

T CELLS

VISTA is a checkpoint regulator for naïve T cell quiescence and peripheral tolerance

Mohamed A. ElTanbouly[1]*, Yanding Zhao[2,3]*, Elizabeth Nowak[1], Jiannan Li[4], Evelien Schaafsma[2,3], Isabelle Le Mercier[5], Sabrina Ceeraz[6], J. Louise Lines[1], Changwei Peng[7,8], Catherine Carriere[9], Xin Huang[9], Maria Day[9], Brent Koehn[10], Sam W. Lee[11], Milagros Silva Morales[7], Kristin A. Hogquist[7,8], Stephen C. Jameson[7,8], Daniel Mueller[7,8], Jay Rothstein[9], Bruce R. Blazar[10,12], Chao Cheng[3,13]†, Randolph J. Noelle[1,9]†

Negative checkpoint regulators (NCRs) temper the T cell immune response to self-antigens and limit the development of autoimmunity. Unlike all other NCRs that are expressed on activated T lymphocytes, V-type immunoglobulin domain-containing suppressor of T cell activation (VISTA) is expressed on naïve T cells. We report an unexpected heterogeneity within the naïve T cell compartment in mice, where loss of VISTA disrupted the major quiescent naïve T cell subset and enhanced self-reactivity. Agonistic VISTA engagement increased T cell tolerance by promoting antigen-induced peripheral T cell deletion. Although a critical player in naïve T cell homeostasis, the ability of VISTA to restrain naïve T cell responses was lost under inflammatory conditions. VISTA is therefore a distinctive NCR of naïve T cells that is critical for steady-state maintenance of quiescence and peripheral tolerance.

Checkpoint regulation of T cell function is governed by coinhibitory molecules (e.g., CTLA-4, VISTA, LAG-3, TIM-3, and TIGIT), which act in concert to fine-tune T cell response and fate (*1*). The importance of these negative checkpoint regulators (NCRs) has been clearly established for cancer and infectious diseases (*2*), but because NCRs are expressed only after T cell activation, it has not yet been determined if they play a role within the naïve T cell compartment to maintain quiescence or response to self-antigen (*1*–*4*). Quiescent T cells make up the overwhelming majority of T lymphocytes in the periphery. Maintaining T cell quiescence and tempering self-reactivity are active processes necessary for survival of an individual. Quiescence regulation is controlled by a diverse set of transcriptional regulators, including forkhead box (FOX) proteins, Kruppel like factors (KLFs), and APRO (Tob1) family members (*5*–*7*). Through control of cellular state and cell cycle arrest, these transcription factors (TFs) reduce the resources necessary to maintain the vast repertoire of resting T cells, of which only an extremely limited frequency will be clonally selected by antigen during the lifetime of the host. Impaired function or deletion of these intracellular mediators can lead to T cell activation and a breakdown in self-tolerance (*2*–*4*, *8*–*10*). Therefore, quiescence and tolerance are functionally linked. Although insights into the intracellular mediators that control naïve T cell quiescence are being realized, the checkpoint regulators expressed on T cells that regulate quiescence are yet to be described.

V-type immunoglobulin domain-containing suppressor of T-cell activation (VISTA) is a member of the B7 family that is distinct from other negative checkpoint molecules in that it is constitutively expressed on naïve T cells. Mice deficient in VISTA show an enhanced frequency of antigen-experienced memory $CD4^+$ $CD44^{hi}$ T cells, heightened cytokine production, and an increased propensity to develop autoimmunity (*11*–*14*). In this regard, genetic deletion of VISTA in the 2D2 myelin oligodendrocyte glycoprotein (MOG)–specific $CD4^+$ T cell receptor (TCR) transgenic (Tg) mouse model of spontaneous autoimmunity results in greatly enhanced inflammatory disease and diminished survival (*13*). Taken together, these observations support the hypothesis that VISTA deficiency results in a breakdown of self-tolerance and the development of inflammatory T cell self-reactive responses. That VISTA is expressed on naïve T cells and lost upon immunization (*12*, *13*) further suggests that its impact on controlling self-tolerance is within the naïve T cell subset.

[1]Department of Microbiology and Immunology, Norris Cotton Cancer Center, Geisel School of Medicine at Dartmouth, Lebanon, NH, USA. [2]Department of Molecular and Systems Biology, Geisel School of Medicine at Dartmouth, Hanover, NH, USA. [3]Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Hanover, NH, USA. [4]Adimab LLC, Lebanon, NH, USA. [5]KSQ Therapeutics, Cambridge, MA, USA. [6]Immunology Discovery, Janssen Research and Development LLC, Spring House, PA, USA. [7]Division of Rheumatic and Autoimmune Diseases, Center for Immunology, University of Minnesota, Minneapolis, MN, USA. [8]The Center for Immunology, University of Minnesota, Minneapolis, MN, USA. [9]ImmuNext Corporation, Lebanon, NH, USA. [10]Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN, USA. [11]Yale University School of Medicine, New Haven, CT, USA. [12]Department of Medicine, University of Minnesota Medical School, Minneapolis, MN, USA. [13]Department of Medicine, Baylor College of Medicine, Houston, TX, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: rjn@dartmouth.edu (R.J.N.); chao.cheng@bcm.edu (C.C.)

Results

VISTA deficiency disrupts the naïve T cell repertoire by reducing quiescence and enhancing T cell activation

VISTA has been shown to act as a coinhibitory receptor on resting $CD4^+$ T cells that negatively regulates T cell activation (*12*, *13*, *15*). VISTA-deficient $CD4^+$ T cells exhibit enhanced proliferation and effector responses to anti-CD3 and antigenic stimulation in vitro (*15*). $VISTA^{-/-}$ mice have heightened antitumor responses to autologous tumors and are more susceptible to death resulting from ConA-induced hepatitis (*12*, *13*, *15*). Although the steady-state percentage of $CD4^+$ T cells was not enhanced in $VISTA^{-/-}$ mice, two groups independently reported an increase in "antigen-experienced" $CD44^{hi}$ $CD62L^{lo}$ $CD4^+$ T cells in the spleens and peripheral blood of $VISTA^{-/-}$ mice (*12*, *13*). Under conditions of conditional VISTA deficiency within the $CD4^+$ T cell compartment, we observed a similar increase in the frequency of antigen-experienced $CD4^+$ T cells, suggesting that the intrinsic loss of VISTA was sufficient for the rise of this activated T cell subset (fig. S1A) (*12*, *13*). That VISTA is expressed on >97% of naïve T cells (fig. S1B) and is lost under inflammatory conditions suggests that its impact on controlling T cell responses is intrinsic to the naïve T cell subset. On the basis of these findings, we interrogated the naïve $CD4^+$ T cell compartment to determine if VISTA altered the steady state to influence their differentiation to antigen-experienced $CD44^{hi}$ cells.

We deleted VISTA in the $CD4^+$ T cell compartment using CD4-Cre mice (hereafter referred to as CD4-Cre-$VISTA^{-/-}$) and performed single-cell RNA-sequencing (scRNA-seq) to examine the role of VISTA in naïve $CD4^+$ T cell transcriptional heterogeneity. scRNA-seq analysis of sorted, naïve ($CD44^{lo}$ $CD62L^{hi}$) $CD4^+$ T cells from CD4-Cre-$VISTA^{-/-}$ mice versus their littermate wild-type (WT) controls revealed a shift in the transcriptional phenotype and heterogeneity within the T cell compartment (Fig. 1A; fig. S1, C and D; and table S1). The most significant phenotypic shift was observed for clusters 1 and 2 (described below). Cluster 1 represents a population of quiescent T cells marked by an up-regulation of *Klf2* and its effectors, which include *Ccr7* (fig. S1 and table S1). This module was reported to be critical for the active maintenance of T cell quiescence and inhibition of proliferation (*7*, *9*, *16*). There is also data supporting the importance of KLF2 in regulating thymocyte trafficking (*17*, *18*). This cluster also included antiproliferative genes such as *Klf6*, which (similar to *Klf2*) up-regulates the negative

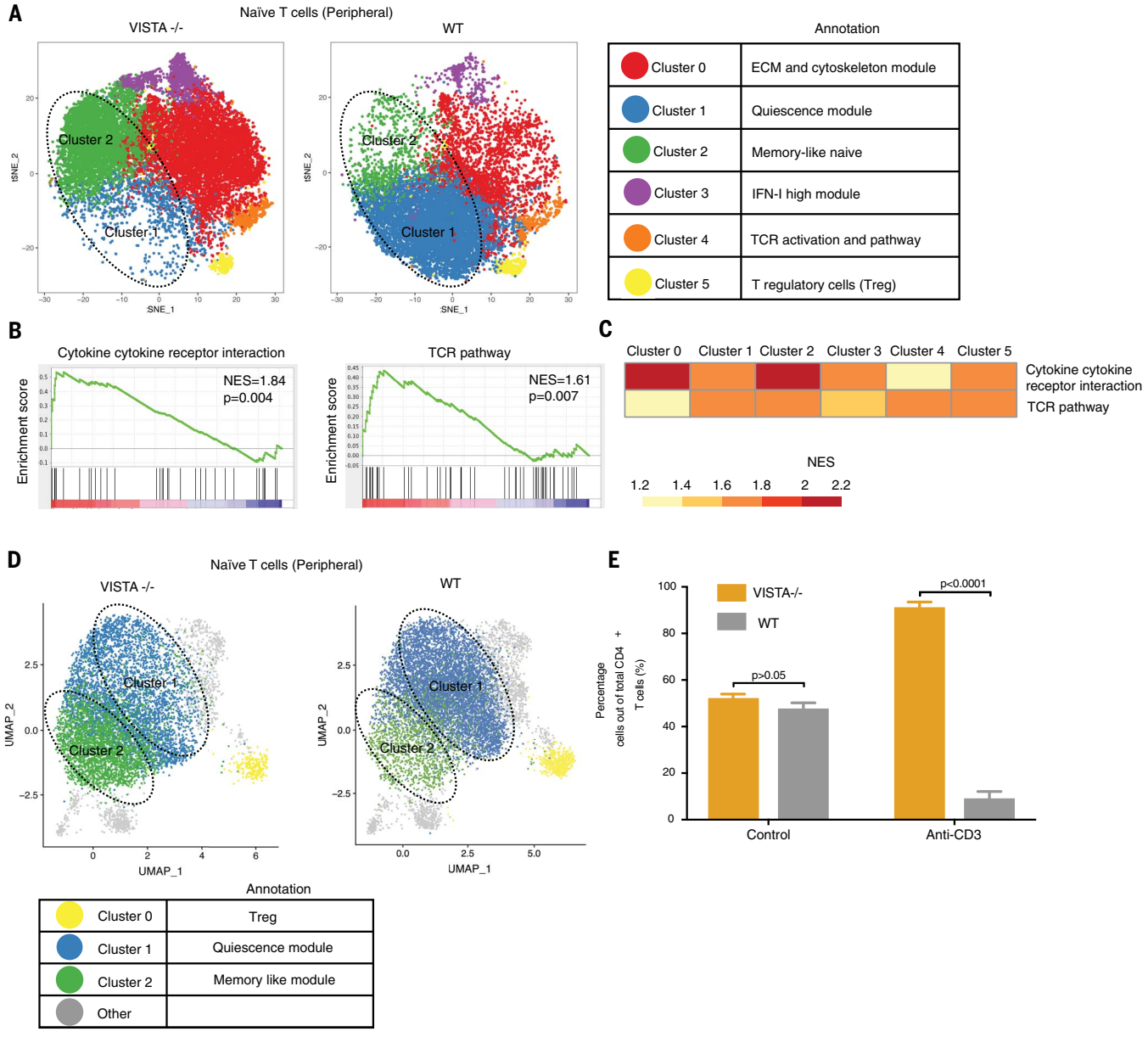


Fig. 1. Intrinsic VISTA deficiency alters heterogeneity in the naïve CD4$^+$ T cell pool. (**A** to **C**) scRNA-seq was performed on naïve CD4$^+$ T cells from WT mice and CD4-Cre-VISTA$^{-/-}$ mice (for which VISTA deficiency is restricted to the CD4$^+$ T cell compartment). (A) t-Distributed stochastic neighbor embedding (t-SNE) plot showing the cluster distribution of FACS-sorted single naïve (>99%) (CD62L^{hi} CD44^{lo}) CD4$^+$ T cells (sorted on the basis of the 20% lowest CD44$^-$ from the negative gate) from CD4-Cre-VISTA$^{-/-}$ and WT littermates. Each dot corresponds to one single cell, colored according to cell cluster. The biological annotation of each cluster is shown in the table on the right. The dashed circles indicate the quiescent T cell cluster (cluster 1) and memory-like naïve T cell cluster (cluster 2). (B) Gene set enrichment analysis (GSEA) pathway enrichment plot indicating the representative gene sets enriched in VISTA$^{-/-}$ versus WT CD4$^+$ T cells. Normalized enrichment score (NES) and *P* values are shown for each gene set. *P* values were calculated by Kolmogorov-Smirnov test. (C) Heatmap showing GSEA analysis as performed in (B) for each cluster. The NES is shown for each gene set across clusters. (**D**) Uniform manifold approximation and projection (UMAP) plot showing the cluster distribution of FACS-sorted single naïve (>99%) (CD62L^{hi} CD44^{lo}) CD4$^+$ T cells from VISTA$^{-/-}$ and WT mice. For (A) to (D), data are representative of two independent experiments with at least three mice per group. (**E**) Ratios of recovered WT versus VISTA$^{-/-}$ CD4$^+$ T cells 5 days after coadoptive transfer into *Rag1*$^{-/-}$ hosts with in vivo anti-CD3 stimulation or control immunoglobulin G (IgG). Data are representative of four independent experiments with at least four mice per group. Each bar indicates the mean value, and each error bar refers to one standard deviation (SD). Student's *t* tests were performed to compare WT with VISTA deficiency (VISTA$^{-/-}$) under each condition (i.e., no treatment versus anti-CD3).

cell cycle regulator p21 (*16*, *19*, *20*). *Btg1* and *Btg2* are members of the Tob gene family, which have critical antiproliferative functions and whose member Tob1 is well described in regulating T cell quiescence and anergy (*21*–*23*). Both genes were defining markers for cluster 1, which is predominant in the WT and lost in the VISTA$^{-/-}$ naïve CD4$^+$ T cell population. VISTA deficiency reduced the abundance of naïve T cells in cluster 1 by more than 10-fold compared with WT CD4$^+$ T cells. Cluster 2, a population augmented by fivefold in VISTA$^{-/-}$ CD4$^+$ naive T cells, was marked by an up-regulation of a stem-cell memory-like program defined by the distinct up-regulation of

Tcf7, *Bcl2*, and *Il7r* (table S1). Of note, one of the defining genes for this cluster was *Klf3*, which has been reported to antagonize KLF2 function (*24*). In addition, this cluster had a higher expression of the costimulatory receptors *Slamf6* and *Ifngr1*, which suggests that these cells are transcriptionally poised for better effector cell responses (*25*–*27*) and that VISTA may play an intrinsic role in maintaining naïve T cell identity and homeostasis. Cluster 0 (enhanced by 2.5-fold by loss of VISTA) was defined by an up-regulation of extracellular matrix interaction pathways and genes such as actin (*Actg1* and *Actb*) and *Cnn2*, which mediate cytoskeletal rearrangements (tables S1 and S2). The role of these pathways has been now appreciated in sustaining the immunological synapse and driving T cell effector function and is supported by pathway analysis of this cluster (*28*, *29*). We observed an up-regulation of TCR pathways in this cluster compared with the remaining population, supporting an overall enhanced abundance of greater TCR signaling in VISTA$^{-/-}$ T cells (table S2).

Two independent groups previously showed that VISTA deficiency or targeting may affect induced regulatory T cell (T_{reg}) induction and stability (*30*, *31*). There were no significant differences in the abundance of the CD44^{lo} T_{reg} cluster (cluster 5) or a cluster defined by enhanced TCR activation–associated transcriptional differences (cluster 4). As stated above, as with the other clusters, there was enhanced TCR signaling imparted by VISTA deficiency (Fig. 1B and tables S2 and S3). In all of the clusters, we observed a significant up-regulation in multiple TCR signaling and cytokine response pathways in VISTA-deficient cells (Fig. 1, B and C, and table S2). Clusters 3 to 5 accounted for less than 7% of the total naïve T cell population and showed no significant differences in abundance between WT and VISTA$^{-/-}$ groups (not discussed in detail).

Given that loss of VISTA reduced quiescent T cells and altered the naïve CD4 T cell repertoire at the gene expression level, we hypothesized that VISTA maintains the epigenetic program for naïve T cell quiescence. We used the assay for transposase-accessible chromatin using sequencing (single-cell ATAC-seq) (*32*) on naïve CD4 T cells from CD4-Cre-VISTA$^{-/-}$ mice or littermate controls. The cell population changes observed were in agreement with those predicted by our scRNA-seq studies. For example, a significant reduction in quiescent T cells (cluster 1) and an increase in memory-phenotype cells (cluster 2) in the VISTA$^{-/-}$ naïve CD4$^+$ T cells were seen (Fig. 1D; fig. S1, E to Q; and table S4). As has been reported by several studies of T cells poised to respond to TCR signaling (*33*, *34*), cluster 2 had an enhancement in the accessibility of multiple TCR effectors (*Nr4a1*, *Cd247*, *Jun*, *Fos*, *Lat*, *Nr4a1*, *Dgka*, and *Nfkb1*) as well as *Cd4*, *Icos*, and *Cd40lg* (Fig. 1D; fig. S1, H to P; and table S4). There is accumulating evidence suggesting that memory cells have significantly greater chromatin accessibility to TCR effector genes (*35*, *36*). Of note, we also observed enhanced accessibility to genes up-regulated in the VISTA$^{-/-}$ memory-phenotype cluster, such as *Tcf7*, *Ifngr1*, *Bcl2*, and *Il7ra*, which supports the suggestion that VISTA deficiency epigenetically primes the naïve T cell repertoire toward a more TCR-responsive memory-like phenotype. Although memory regulators such as *Lef1*, *Zbtb20*, and *Runx3* were not differentially expressed between the quiescent and memory-like clusters at the mRNA level, they had greater chromatin accessibility in the memory-like cluster cells enhanced by VISTA deficiency (table S4). On the other hand, the quiescent cluster representing the majority of WT naïve CD4$^+$ T cells was defined by enhanced chromatin accessibility to *Klf2*, *Klf6*, *Btg1*, and *Btg2*, which are the defining markers of quiescent cells by scRNA-seq analysis (Fig. 1, A and D; fig. S1, H, K, and L; and table S4). In addition, this cluster also had an enhancement in the accessibility of important quiescence factors like *Foxp1*, *Foxo-1*, and *Runx1* (*37*–*39*) (table S4).

One alternative possibility was that the reduced quiescence state of cluster 2 and its expansion in the absence of VISTA were a consequence of increased autoreactive TCR repertoires in this cluster. To address this hypothesis, we performed single-cell TCR sequencing paired with gene expression to resolve the TCR sequences of each cell per cluster. In addition, we extensively reviewed the literature and manually curated a database of more than 50 TCR Vβ CDR3 sequences from CD4$^+$ T cells in multiple models of autoimmunity matching different autoantigen specificities because certain Vβ genes were associated with autoreactivity (*40*) (fig. S2A). We found less than 100 cells across all clusters (out of 40,000 total cells) in the naïve CD4 T cell repertoire that matched these TCR sequences (*40*–*42*). Because this analysis was not sufficient to capture the landscape of autoreactivity in the naïve CD4$^+$ T cell population, we performed single-cell TCR sequencing on CD44^{hi} CD4$^+$ T cells (10× genomics, paired α and β chain) because these cells have reacted to self-antigen in an unimmunized mouse (fig. S2B) (*43*). In this analysis, only the Vβ TCR repertoire with full length and productivity were chosen. We performed CDR3-region sequence alignment (*44*) and chose the Vβ CDR3 sequences that were fully matched between CD44^{lo} and CD44^{hi} cells. We identified a total 4971 "potentially" autoreactive CD44^{lo} cells with 1606 unique TCRs. Then, we quantified the fraction of autoreactive T cells in each CD44^{lo} cluster (clusters having less than 1000 cells were excluded). The fraction of autoreactive T cells was almost evenly distributed among all CD44^{lo} clusters in both VISTA$^{-/-}$ and WT mice (fig. S2B). Because CD44^{hi} CD4$^+$ T cells in an unimmunized mouse may not represent the prototypical autoreactive repertoire seen in autoimmune disease, we performed single-cell TCR sequencing on CD4$^+$ T cells sorted by fluorescence-activated cell sorting (FACS) from B6 Fas lpr mice, an established lupus model for which autoreactive T cells have been reported (*45*). We also performed the same sequencing procedure on CD4$^+$ T cells from Bim-deficient mice, which fail to delete autoreactive T cells during negative selection (*46*). This allowed us to generate full CDR3 sequences for autoreactive TCRs from two independent models of autoimmunity. In both TCR sequence datasets, there was an overlap in the TCR sequences (around 4 to 5%) for each of the naïve CD4 T cell clusters (fig. S2, C to F). However, there was no impact of VISTA loss on the percentage or distribution of the autoreactive T cells. Our interpretation is that the changes in the clusters imposed by the loss of VISTA are not due to changes in TCR specificity of the constituting cells but rather mostly due to a change in the cell state.

To determine if the phenotypic changes in the naïve T cell repertoire were imparted by VISTA deficiency at the mature T cell stage in the periphery or were a consequence of a potential role that VISTA plays in thymocyte development, analysis of the impact of VISTA deficiency on thymocyte heterogeneity was studied. VISTA is constitutively expressed on naive CD4$^+$ T cells and also on single-positive thymocytes (fig. S3, A and B). Flow cytometric analysis of thymocyte subset percentages did not show any impact of intrinsic VISTA deficiency on the thymocyte numbers or frequency (fig. S3B). In addition, scRNA-seq of thymocytes from VISTA$^{-/-}$ or WT littermates did not reveal any differences in heterogeneity of the thymic repertoire (fig. S3, C to E, and table S5). Analysis of the CD4$^+$ lineage differentiation trajectory from double-positive thymocytes to the naïve peripheral T cell stage using the well-established Monocle algorithm (*47*) did not elucidate an impact of VISTA deficiency on the route of CD4$^+$ T cell differentiation (fig. S3, G and H). This suggests that VISTA deficiency did not alter the differentiation route of thymocytes to mature T cells and exclusively exerted an impact on naïve T cell fate in the peripheral compartment.

A series of experiments was performed to gain insights into whether expression of KLF2 was correlated with the expression of other quiescence factors in T cells and, in addition, correlated with VISTA expression. The data presented show that VISTA deficiency results in a global reduction of quiescence regulators such as *Klf2*, *Klf6*, *Gimap5*, and Tob gene family members *Btg1* and *Btg2* (*9*, *16*, *19*–*23*, *48*).

Therefore, there was sufficient evidence to suggest that VISTA is necessary for the expression of multiple quiescence regulators. To more directly address if KLF2 expression was coregulated with the other quiescence factors, a KLF2 reporter mouse (*18*) was used. Using this system, we sought to examine whether higher KLF2 expression on naïve $CD4^+$ T cells recapitulated the naïve T cell quiescence phenotype (cluster 1). $KLF2^{hi}$ and $KLF2^{lo}$ naïve $CD4^+$ T cells were electronically cell sorted (on the basis of the 20% highest and lowest expression) and subjected to deep RNA-seq analysis. $KLF2^{hi}$ $CD4^+$ T cells were highly enriched for genes that define the quiescence cluster of naïve T cells (cluster 1), closely mirroring their profile with regard to differential gene expression (fig. S4A). Therefore, higher expression of $KLF2^{hi}$ is correlated with the heightened expression of other quiescence factors. $KLF2^{hi}$ $CD4^+$ T cells additionally up-regulated several established quiescence regulators such as *Tob1*, *Foxp1*, *Foxo1*, and *Tgfbr2* (fig. S4B) (*6*, *21*). Next, we examined the relationship between VISTA and KLF2 expression, a defining cluster 1 TF and marker. As such, flow cytometric analysis revealed a strong direct correlation between VISTA and KLF2 expression, because increased KLF2 expression ($KLF2^{hi}$) also showed higher VISTA expression (fig. S4, C and D). Of note, $KLF2^{hi}$ $CD4^+$ had significantly higher VISTA mRNA expression. The same RNA-seq analysis was conducted for $VISTA^{hi}$ versus $VISTA^{lo}$ naïve T cells, and $VISTA^{hi}$ $CD4^+$ T cells showed greater KLF2 (as well as *Klf6*, *Btg1*, and *Btg2*). This bidirectional relationship (Pearson correlation coefficient of 0.87) between KLF2 and VISTA is presented as a correlation plot (fig. S4E). Taken together, these data provide compelling evidence using flow cytometry and RNA-seq that VISTA regulates KLF2, an important TF with roles in T cell quiescence. Given the relationship between VISTA and KLF2, we posit that enhanced VISTA expression on the naïve T cell compartment ($CD44^{lo}$ $CD62L^{hi}$) correlates with greater quiescence and the naïve phenotype (fig. S4F). We sought to track the impact of graded VISTA expression on naïve T cells. Using deep RNA-seq, we found that $VISTA^{hi}$ naïve T cells display a more quiescent T cell state than the $VISTA^{lo}$ or $VISTA^{-/-}$ T cells at the global gene expression level (fig. S4G). In summary, there was significant up-regulation of *Klf2*, *Klf6*, and *Slfn2* and a dramatic up-regulation of *Foxp1* in $VISTA^{hi}$ T cells, all critical effectors of T cell quiescence (*8*, *37*). However, $VISTA^{lo}$ T cells expressed higher levels of *Nr4a1*, *Myc* (inhibited by *Klf2*) (*9*), *Pdcd1*, *Ctla4*, *Cd5*, *Cd6*, *Cd2*, *Nfkb1*, *Lck*, and *Nfatc1*, all indicative of enhanced TCR signaling, enhanced activity, and reduced quiescence.

Because the percentage of $CD44^{hi}$ memory-phenotype (MP) $CD4^+$ T cells is enhanced in steady-state unimmunized $VISTA^{-/-}$ mice (fig. S1A) and VISTA deficiency skews the naïve $CD4^+$ T cells toward a less quiescent memory-like phenotype at both the transcriptional and epigenetic levels (Fig. 1 and fig. S1), we investigated how VISTA-deficiency influenced the expansion of the naïve $CD4^+$ T cells toward $CD44^{hi}$ MP cells by scRNA-seq profiling of this population (fig. S5). At least for the $CD4^+$ T cell lineage, conversion of naïve ($CD62L^{hi}$ $CD44^{lo}$) T cells to $CD44^{hi}$ MP cells requires antigen encounter (self but not commensal antigen) and sufficient TCR stimulation (*43*, *49*). Examination of the VISTA-deficient $CD44^{hi}$ MP T cells revealed that the most dramatic enhancement caused by VISTA deficiency was a more than threefold increase in T helper 1 cell (T_H1) effector phenotype cells (cluster 1) (fig. S5, A to C). Cells in cluster 1 up-regulate the T_H1 master TF *Tbx21* (T-bet) as well as the characteristic T_H1 effector molecules *Ifng*, *Ccl5*, and *Cxcr3* (table S6) (*50*). Globally, there was an up-regulation of effector molecules *Ifng*, *Ccl5*, and *Cxcr3*, in addition to costimulatory molecules such as *Cd7*, *Cd40lg*, *Cd69*, and *Ly6c* (fig. S5 and table S6). On the other hand, there was more than a threefold reduction of a coinhibitory module group of cells (cluster 5) in the antigen-experienced repertoire defined by up-regulation of multiple checkpoint regulators such as PD-1, LAG-3, TIGIT, CD73 (*Nt5e*), FR4 (*Izumo1r*), BTLA, and Nrp1 and other regulators of T cell dysfunction such as *c-Maf*, *NFATc*, *Tox*, and *Tox2* (fig. S5E) (*51*). Indeed, analysis of the whole T cell population supported this because $VISTA^{-/-}$ had markedly reduced expression of the coinhibitory regulators (fig. S5F). At the cell-state level, we observed that CD4-Cre-$VISTA^{-/-}$ MP $CD4^+$ T cells up-regulate greater downstream TCR activation genes as marked by significant global up-regulation of the AP-1 and JNK TF network (*Jun*, *Junb*, *Jund*, and *Fos*) as well as *Cd69*, *Nr4a1* (*Nur77*), and nuclear factor κB (NFκB) pathway effectors (fig. S5D). The majority of these genes had a greater chromatin accessibility in the $VISTA^{-/-}$ naïve $CD44^{lo}$ cells (fig. S1 and table S4). Heightened TCR signaling in $VISTA^{-/-}$ cells and the heightened expression of TCR activation genes and other markers of T cell activation would be predicted by the work by Paul and colleagues who showed that TCR antigen encounter and costimulation are essential for establishment of MP $CD4^+$ T cells in unimmunized mice and that this population expresses high levels of T-bet and interferon-γ (IFN-γ) (*49*). It is also in agreement with reports that stronger TCR signals favor T_H1 polarization (*52*) and that VISTA-deficiency on naïve T cells enhances production of IFN-γ and other T_H1 cytokines upon TCR stimulation in vitro (*12*). Therefore, intrinsic VISTA expression on the naïve T cell is necessary for restraining T cell activation and maintaining quiescence, and VISTA-deficiency engenders a T_H1 proinflammatory phenotype in the absence of the appropriate costimulatory or immunizing signals.

Loss of quiescence has been repeatedly correlated with reduced tolerance susceptibility of T cells (*7*). Given the enhanced activation phenotype of the $VISTA^{-/-}$ T cells and the reduction in the quiescent cluster, we hypothesized that peripheral, TCR-induced, deletional tolerance of $VISTA^{-/-}$ T cells may be impaired. To this end, we used a coadoptive transfer system whereby naïve $VISTA^{-/-}$ and WT $CD4^+$ T cells are transferred at equal ratios into T cell–deficient $Rag^{-/-}$ hosts and administered anti-CD3 monoclonal antibody (mAb) or control mAb (Fig. 1E). It has been shown that under these conditions, TCR signaling by anti-CD3 induces the deletion of T cells in vivo (*53*, *54*). In the absence of anti-CD3 stimulation, we could not detect a marked difference between WT and $VISTA^{-/-}$ $CD4^+$ T cell numbers, indicating no notable advantage of VISTA deficiency on homeostatic T cell expansion or survival (Fig. 1E and fig. S4G). Only when anti-CD3 stimulation was provided could we detect a marked enhancement in the recovered numbers of $VISTA^{-/-}$ $CD4^+$ T cells compared with the number of WT $CD4^+$ T cells (Fig. 1E and fig. S5G). This supports the scRNA-seq data on the naïve T cell phenotype, establishing that VISTA deficiency leads to a loss in quiescence and a reduced susceptibility to TCR-induced deletion. T cells from CD4-Cre × $VISTA^{fl/fl}$ mice also recapitulated this phenotype, establishing that the resistance to anti-CD3 deletion was T cell intrinsic (fig. S5H). These data suggest that loss of quiescence affects the fate of TCR-triggered T cells in vivo.

Agonistic anti-VISTA mAbs enhance TCR-dependent peripheral T cell deletion

Given that we observed the functional impact of VISTA deficiency to be a reduced susceptibility to anti-CD3–induced deletion, we hypothesized that antibody-based activation of VISTA would enhance TCR-induced T cell deletion. Chen and colleagues introduced a class of anti-VISTA "agonists" and showed in multiple systems, including graft-versus-host disease (GVHD), that the agonist antibodies suppressed T cell immune responses [reviewed in (*55*)]. We developed both anti-mouse–specific (clone 8G8) and anti-human–specific (clone 803) VISTA agonists to assess the impact of VISTA engagement on T cell fate on TCR engagement. The isotype and functional properties of the anti-VISTA clones used in this study are detailed in table S7 (fig. S6, A to E). These mAbs are specific for mouse or human VISTA, and both suppress GVHD. In addition, mouse anti-VISTA (anti-mVISTA) (8G8) has demonstrated immunosuppressive properties in multiple murine models of inflammation

and autoimmunity (fig. S6, B to E). To assess the impact of anti-VISTA agonist on tolerogen-induced T cell deletion, naïve OVA-specific transgenic $CD4^+$ T cells (OT-II) were adoptively transferred to antigen-bearing hosts (Act-Ova) or antigen-deficient B6 controls and treated with anti-mVISTA (8G8) (*56*, *57*) (Fig. 2A and fig. S6F). This system provides a TCR engagement signal (OVA) but no overt inflammatory signal, which are both known to promote tolerance induction (*58*, *59*). Administration of anti-VISTA overwhelmingly reduced the frequency of OT-II T cells in the Act-Ova–expressing hosts but not in B6 hosts. Furthermore, there was an enhanced percentage of dead OT-II T cells, suggesting that anti-VISTA enhanced tolerogenic T cell death (Fig. 2B and fig. S6, G and H). VISTA has been reported to participate in the uptake and clearance of apoptotic cells (*14*). We therefore investigated whether the change in percentage of dead cells may be due to a role for agonistic anti-VISTA (8G8) on dead cell clearance of OT-II cells. The impact of anti-VISTA on the uptake of apoptotic thymocytes by macrophages was assessed by flow cytometry (fig. S6I). Although VISTA-deficiency had a significant impact on the uptake of apoptotic cells (*14*), anti-VISTA (8G8) did not demonstrate any significant inhibitory activity. Therefore, the enhanced death of OT-II T cells observed in 8G8-treated Act-OVA mice is likely due to enhanced cell death caused by augmented deletional tolerance with no impact on clearance. It remained possible that anti-VISTA may have enhanced antigen-induced T cell tolerance by direct killing of other $VISTA^+$ immune populations, most prominently antigen-presenting cells. However, examination of the abundance of various immune populations after anti-VISTA treatment revealed no significant impact of the antibody on their numbers or frequency (fig. S6J). There were no significant reductions in the frequency T cell or various myeloid populations upon treatment of mice with either VISTA agonist or antagonist antibodies under steady-state conditions (fig. S6J). These populations include $CD11b^+$ myeloid cells, neutrophils, monocytes, and dendritic cells.

To prove that VISTA-induced T cell loss was due to a direct effect of anti-VISTA antibody binding to T cells, OT-II T cells expressing human VISTA (hVISTA) were specifically targeted with an anti-hVISTA antibody and transferred into WT host mice. hVISTA-expressing OT-II cells were obtained subsequent to interbreeding with hVISTA knock-in (KI) mice. Extensive validation of hVISTA lineage and expression levels and the specificity and affinity of the anti-hVISTA mAb (803) in hVISTA KI mice were evaluated (fig. S7 and materials and methods). Flow cytometric analysis of hVISTA expression on myeloid and T cell subsets in hVISTA KI mice revealed similar levels of expression to murine VISTA in WT mice (fig. S7, A to C). The specificity of anti-hVISTA (803) was validated in multiple systems (fig. S7, B and C) and also in a Jurkat cell line transduced to express hVISTA versus WT Jurkat cells (fig. S7D). In contrast to VISTA deficiency, anti-hVISTA induced profound reductions in hVISTA $CD4^+$ T cells with anti-CD3 tolerization (Fig. 1E and fig. S7E). Similar to the anti-mVISTA clone 8G8, anti-hVISTA (803) reduced the frequency of adoptively transferred OT-II T cells expressing hVISTA upon administration of soluble OVA peptide but not in the absence of peptide (fig. S7F). These findings show that targeting hVISTA exclusively on the T cell surface, together with TCR engagement, results in a selective reduction of targeted cells. As has been observed with anti-mVISTA (8G8), there was no impact of anti-hVISTA (803) on the abundance of the different immune lineages (fig. S7G). To confirm that the augmented T cell tolerance induced by agonistic anti-VISTA and antigen was mediated by targeting the donor antigen-specific $CD4^+$ T cells, and not due to potential indirect effects of targeting the $VISTA^+$ myeloid cells, we adoptively transferred hVISTA OT-II cells in the presence of OVA and exclusively targeted the host using anti-mVISTA (8G8), sparing the donor hVISTA OT-II cells. In this case, there was no impact on donor antigen-specific T cell numbers. These findings indicate that targeting the T cell compartment is necessary and sufficient for the augmented tolerance by anti-VISTA agonists (fig. S7H).

VISTA regulates the fate of tolerized, endogenous antigen-specific T cells

Our findings suggest that VISTA regulates the fate of TCR-engaged T cells in vivo. To rigorously test this hypothesis, we studied the impact of VISTA targeting on the fate of endogenous, antigen-specific $CD4^+$ T cells under tolerogenic conditions using soluble peptide-loaded major histocompatibility complex (pMHC)–based tetramer enrichment systems (*60*–*62*). Analysis of tetramer-positive $CD4^+$ T cells revealed that upon the administration of soluble 2w1s antigen, VISTA deficiency enhanced the number of 2w1s:$I\text{-}A^b$–specific $CD4^+$ T cells by more than twofold under conditions of tolerance induction (Fig. 2C). We tested whether VISTA blockade would recapitulate the outcome observed with VISTA deficiency on 2w1s:$I\text{-}A^b$ response to antigen under tolerogenic conditions. To do this, we used the anti-VISTA antagonist antibody (13F3), a well-established VISTA-blocking clone we previously reported (*30*, *63*). Like VISTA deficiency, anti-VISTA blockade increased the number of 2W1s:$I\text{-}A^b$ specific T cells upon tolerogenic peptide administration (Fig. 2E). By contrast, agonistic anti-VISTA imparted about a twofold reduction in the number of 2w1s: $I\text{-}A^b$–specific $CD4^+$ T cells under the same tolerogenic conditions (Fig. 2F). The opposing impacts of VISTA deficiency and anti-VISTA agonist on T cell tolerization were also observed using endogenous $CD4^+$ MOG:$I\text{-}A^b$ endogenous T cells under conditions of MOG peptide administration (fig. S8, A and B). All studies thus far evaluated the impact of anti-VISTA on T cells under conditions of exclusive TCR engagement and in the absence of inflammation. When mice were immunized with 2w1s peptide and lipopolysaccharide (LPS), there was no impact of VISTA deficiency on the number of endogenous 2W1s-specific T cells (Fig. 2D). Similarly, agonistic anti-VISTA failed to impart a significant impact on antigen-specific T cells under inflammatory conditions (Fig. 2G). This presents evidence that VISTA engagement is important for restraining T cell expansion under tolerogenic conditions and that inflammation can supersede the impact of VISTA on T cell fate.

In addition to deletion, peripheral $CD4^+$ T cell tolerance is regulated by multiple other mechanisms. Therefore, we investigated whether VISTA affected the emergence of antigen-specific anergic and/or regulatory antigen-specific $CD4^+$ T cells. T cell activation under conditions that lack costimulation induced a state of hyporesponsiveness marked by proliferation arrest and markedly diminished effector cytokine production upon restimulation (*64*). It was previously reported that anergic $CD4^+$ T cells up-regulated the two surface markers CD73 (*Nt5e*) and FR4 (*Izumo1r*) (*65*, *66*). Indeed, single-cell analysis of total $CD44^{hi}$ $CD4^+$ T cells from unimmunized mice validates the existence of this naturally anergic $CD4^+$ ($CD44^{hi}$ $Foxp3^-$) T cell population and also identifies multiple additional regulators that participate in T cell anergy such as *NFATc1* and *Nrp1* (fig. S9A). This population, as do the majority of $CD4^+$ T cells, express significant levels of VISTA, so VISTA was therefore not a defining marker for this cluster. To our knowledge, this presents the first full transcriptional profile of naturally anergic $CD4^+$ T cells. We investigated whether agonistic anti-VISTA would reduce antigen-specific $CD4^+$ T cell numbers under tolerogenic conditions by enhancing the number of anergic cells. Analysis of the percentage of anergic 2w1s:$I\text{-}A^b$–specific $CD4^+$ T cells by $CD73^{hi}$ $FR4^{hi}$ $Foxp3^-$ staining did not show a differential impact of anti-VISTA treatment (fig. S9, B to D). As expected from the phenotypic analysis, anti-VISTA also did not enhance the percentage of anergic cells by cytokine-responsiveness [interleukin-2 (IL-2) and IFN-γ] to in vitro restimulation with phorbol 12-myristate 13-acetate (PMA) and ionomycin (fig. S9E). Of note, we did report similar percentages of anergic antigen-specific $CD4^+$ T cells to those previously published for antigen-induced tolerance of the 2w1s-specific repertoire (*65*).

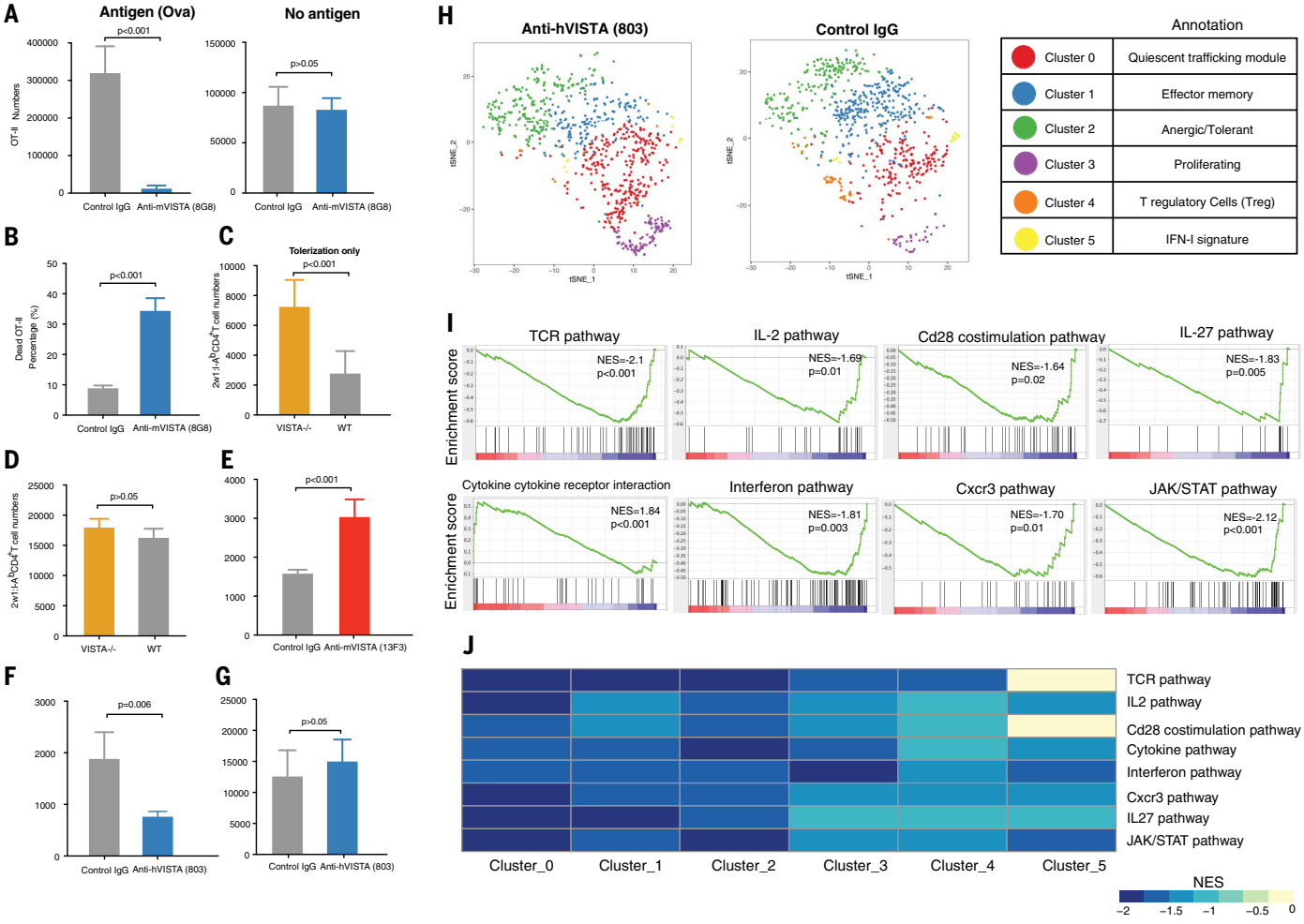


Fig. 2. Agonistic anti-VISTA antibodies augment T cell tolerance, which is abrogated by VISTA deficiency. (**A**) Recovered numbers of OT-II CD4$^+$ T cells in the spleen of anti-mVISTA (8G8)–treated or hamster IgG control–treated mice transferred into Act-Ova (left) or B6 hosts (right) 48 hours after adoptive transfer. (**B**) Percentage of dead OT-II CD4$^+$ T cells out of total recovered OT-II cells from anti-mVISTA (8G8)–treated or hamster IgG control–treated mice. Data are representative of two independent experiments with five mice per group. (**C**) Two doses of 2w1s peptide (100 μg) were intravenously injected into WT or VISTA$^{-/-}$ mice on days 0 and 3, respectively, followed by tetramer enrichment and 2w1s:I-A^b cell number quantification on day 7. (**D**) 2w1s peptide (100 μg) with 5 μg of LPS was intravenously injected into WT or VISTA$^{-/-}$ mice on day 0, followed by tetramer enrichment and 2w1s:I-A^b cell number quantification on day 7. (**E**) Two doses of 2w1s peptide were intravenously injected into antagonistic anti-mVISTA (13F3)–treated or hamster IgG control–treated mice on days 0 and 3, respectively, followed by tetramer enrichment and 2w1s:I-A^b cell number quantification on day 7. (**F**) Two doses of 2w1s peptide were intravenously injected into agonistic anti-hVISTA (803)–treated or isotype control–treated mice on days 0 and 3, respectively, followed by tetramer enrichment and 2w1s:I-A^b cell number quantification on day 7. Data are representative of four independent experiments with at least eight mice per group [(C) to (F)]. (**G**) 2w1s peptide (100 μg) with 5 μg of LPS was intravenously injected into control IgG–treated or agonistic anti-hVISTA (803)–treated mice on day 0, followed by tetramer enrichment and 2w1s:I-A^b cell number quantification on day 7. Data are representative of two independent experiments with eight mice per group. (**H**) t-SNE plot showing the cluster distribution of 2w1:I-A^b peptide–induced CD4$^+$ T cells from agonistic anti-hVISTA (803)–treated and control IgG–treated mice. Each dot corresponds to one single cell, colored according to cell cluster. Biological annotation of each cluster is shown in the table on the right. (**I**) GSEA pathway enrichment plot indicating the representative gene sets depleted in agonistic anti-hVISTA (803)–treated versus isotype control–treated mice. NESs and *P* values are shown for each gene set. *P* values were calculated by Kolmogorov-Smirnov test. (**J**) Heatmap showing GSEA analysis as performed (I) for each cluster. NESs are shown for each gene set across clusters. Sequencing data are representative of two independent repeats with at least two samples from pooled mice per group. For all bar plots [(A) to (G)], each bar indicates the mean value and each error bar refers to one SD; *P* values were calculated by Student's *t* test.

An additional mechanism of peripheral tolerance is through the emergence of antigen-specific Foxp3$^+$ T$_{regs}$ and the inhibition of effector T cell expansion and function (*67*). We asked whether anti-VISTA would change the number of antigen-specific 2w1s: I-A^b–specific Foxp3$^+$ CD4$^+$ T$_{regs}$ and thereby suppress T cell expansion. Analysis of the percentage of Foxp3$^+$ CD4$^+$ T$_{regs}$ did not show any impact of anti-VISTA (fig. S9D). These results suggest that VISTA engagement or blockade did not overtly change anergy induction under tolerogenic conditions.

We performed high-resolution scRNA-seq gene expression profiling to assess the impact of anti-VISTA treatment on tolerized 2w1s:I-A^b–specific CD4$^+$ T cells, which revealed insights into repertoire heterogeneity and cell state at the single-cell level (Fig. 2H and fig. S10A). Surprisingly, there was no significant impact of VISTA agonistic targeting or deficiency on the heterogeneity of the antigen-specific repertoire (Fig. 2H; fig. S10, A and B; and table S8). However, this analysis yielded

two important observations. First, anti-VISTA reduced T cell clonal expansion of the 2w1s:I-A^b repertoire in all clusters (fig. S10C). Second, analysis of pathway activity revealed that VISTA triggering resulted in a global reduction (>80% of the repertoire) in TCR signaling pathways, such as CD28 costimulation, CXCR3 signaling, and cytokine interactions, suggesting a major impact on the global state of tolerized T cells (Fig. 2, I and J). On the other hand, VISTA deficiency promoted an up-regulation of proliferation pathways and globally enhanced cellular transcription and translation (tables S9 and S10). This indicates that VISTA engagement under tolerogenic conditions imparts an immunosuppressive phenotype to augment T cell tolerance in addition to reducing tolerized T cell numbers, which is in support of data with transgenic systems (Fig. 2, A and B, and figs. S6G and S7F).

Sustained expression of VISTA under tolerogenic, but not inflammatory, conditions

Our data suggest that VISTA engagement renders naïve T cells more susceptible to antigen-induced death and down-regulates pathways of TCR signaling. This may support the argument that sustained VISTA expression would prohibit T cell activation. In addition, we found that VISTA engagement on naïve T cells is abolished under inflammatory conditions (Fig. 2, D and G). We therefore investigated whether TCR engagement under inflammatory (antigen with LPS) versus tolerogenic (antigen only) conditions affected VISTA expression on endogenous antigen-specific $CD4^+$ T cells using scRNA-seq of 2w1s:I-A^b–specific $CD4^+$ T cells. In the inflammation setting, we observed a global transcriptional down-regulation of VISTA in tetramer$^+$ cells (Fig. 3, A and B, and table S11), which was supported by flow cytometric analysis (Fig. 3C). There was no change in VISTA expression on total $CD4^+$ T cells, suggesting that the impact was only on antigen-specific T cells (Fig. 3D). As expected, pathways of proliferation, CD28 costimulation, and antigen response were significantly up-regulated under inflammatory conditions (table S12). Furthermore, a major cluster of cells with a tolerant transcriptional phenotype (cluster 3) was exclusive to tolerization and one of the defining markers for this cluster was VISTA (Fig. 3, A and E, and table S11). This cluster included known regulators of T cell suppression and dysfunction such as FR4 (*Izumo1r*), LAG-3 (*Lag3*), BTLA (*Btla*), SHP-2 (*Ptpn11*), Neuropilin-1 (*Nrp1*), *Slfn2*, and *Nr4a1* (*1*, *65*, *68*). These molecules were expressed in addition to tumor necrosis factor receptor superfamily molecules, which mirrors the profile of T cell dysfunction previously reported (*69*). This suggests a potential consequence of VISTA expression under conditions of tolerance but minimal consequence under productive costimulation of T cells.

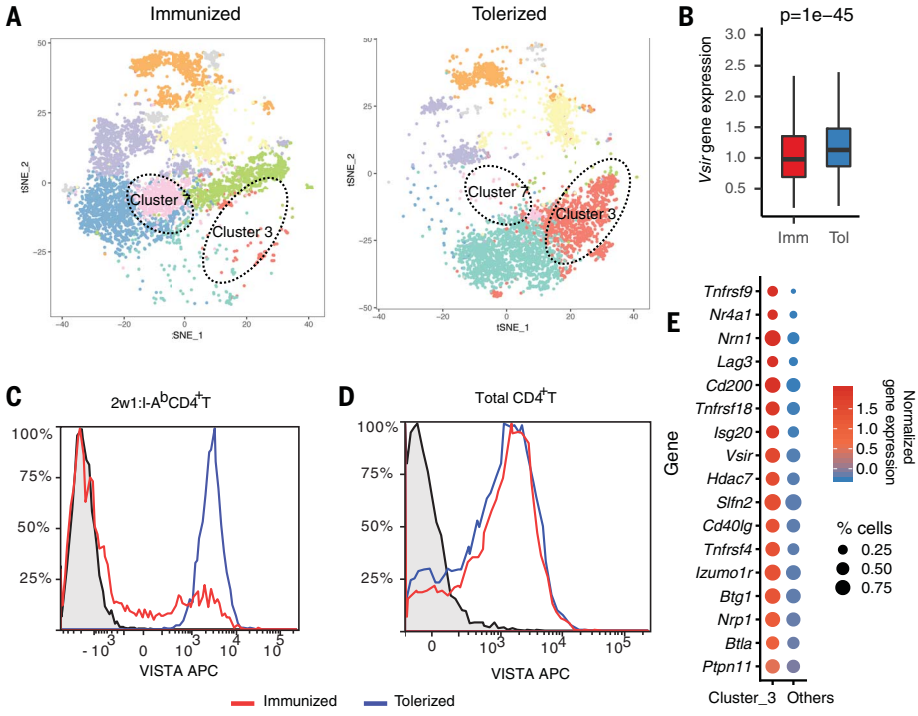


Fig. 3. VISTA expression is reduced under inflammatory, but not tolerogenic, conditions in vivo. To provide a model of antigen-specific stimulation for the antigen-specific $CD4^+$ T cell repertoire studies in Fig. 2, C57BL/6 mice were either immunized using 2w1s peptide with LPS or tolerized using 2w1s peptide only. Then, the phenotype of 2w1s:I-A^b–specific $CD4^+$ T cells was analyzed using scRNA-seq and flow cytometry. (**A**) The t-SNE plot shows the cluster distribution of 2w1:I-A^b peptide–induced $CD4^+$ T cells from immunized (2w1s peptide with LPS) and tolerized (2w1s peptide) mice. Each dot corresponds to one single cell, colored according to cell cluster. The dashed circles indicate the anergic T cell cluster (cluster 3) and T follicular helper T cell cluster (cluster 7). (**B**) Boxplot depicting the difference in *Vsir* gene, which encodes VISTA, expression between immunized and tolerized $CD4^+$ T cells across all clusters. The center line refers to the median value for *Vsir* gene expression. The whisker indicates the 25th to the 75th percentile of *Vsir* gene expression. *P* values were calculated by Wilcoxon rank sum test. (**C** and **D**) Flow cytometric analysis of 2w1s:I-A^b–specific $CD4^+$ T cells (C) or total $CD4^+$ T cells (D) under the same tolerization versus immunization conditions presented in (A). The black plot line and shaded region indicate staining of VISTA knockout $CD4^+$ T cells as a biological control. APC, allophycocyanin. (**E**) Bubble plot showing the average Z-transformed normalized expression of coinhibitory module genes in cluster 3 versus other clusters. The size of each bubble indicates the fraction of cells expressing the represented gene. Data are representative of two independent experiments with at least three mice per group [(A) to (E)].

VISTA targeting induces systemic tolerance and T cell deletion

That the targeting of VISTA with anti-VISTA mAbs at the time of donor T cell transfer can ablate the development of GVHD supports the hypothesis that anti-VISTA agonism can induce antigen-specific T cell tolerance (*12*, *15*, *70*). We confirmed and expanded this data with both agonistic anti-mouse (8G8) and anti-human (803) VISTA clones (Fig. 4, A and B). Agonistic targeting of VISTA exclusively on the donor T cells arrested the development of GVHD (Fig. 4B). Of note, anti-VISTA blockade (13F3) did not affect alloreactive T cell responses or mouse survival, clearly distinguishing the activities of different anti-VISTA mAbs (Fig. 4A). Under the same GVHD conditions, we investigated the fate of alloreactive $CD4^+$ T cells targeted with agonistic anti-VISTA by using the TEa TCR transgenic $CD4^+$ T cells which recognize I-Eα (residues 52 to 68) peptide in the context of I-A^b (*71*). Similar to experiments presented in Fig. 2 (OT-II), after 24 hours, we noted a significant (>75%) reduction in the number of TEa cells when transferred to anti-VISTA–treated, antigen-bearing F1 hosts. However, no reduction was observed when transferred into B6 mice, indicating that both TCR and VISTA engagement were required for VISTA-mediated deletion (Fig. 4C). The loss of TEa $CD4^+$ T cells mediated by antigen and anti-VISTA was not due to the altered localization of T cells in other tissues (fig. S11, A and B). In support of these in vivo observations, scRNA-seq analysis revealed a marked up-regulation of GVHD pathway mediators in VISTA$^{-/-}$ T cells, which were subsequently down-regulated using agonistic anti-VISTA treatment (Fig. 4D). We developed a VISTA deficiency–associated gene signature to reflect

intrinsic VISTA-induced pathway activity change in T cells (see materials and methods and table S13). This signature was validated using bulk RNA sequencing of TCR transgenic VISTA$^{-/-}$ T cells (Fig. 4E) (*13*). The VISTA module was then applied to a well-documented dataset of CD4$^+$ T cell exhaustion versus activation. In this setting, the activated T cells presented a notably higher VISTA module score compared with exhausted T cells, indicating that VISTA$^{-/-}$ related pathways were up-regulated in the activated T cells (Fig. 4F). In addition, VISTA deficiency was also reported to exacerbate autoimmune murine lupus (*11*, *14*). Indeed, peripheral T cells from systemic lupus erythematosus (SLE) patients from two independent datasets presented a higher VISTA module score compared with that of healthy donors (Fig. 4G and fig. S12A). Similarly, peripheral T cells from rheumatoid arthritis (RA) patients presented an even higher VISTA module score (fig. S12B). This evidence supports a broad regulatory role for VISTA in suppressing T cell self-reactivity and autoimmune manifestations. It may also suggest that VISTA could represent a potential diagnostic biomarker for such inflammatory diseases.

Concluding remarks

We report a distinct role for VISTA as a negative checkpoint that regulates naïve T cell quiescence and optimal peripheral T cell tolerance. The genetic loss of VISTA in T cells markedly altered the cell state and heterogeneity of mature naïve T cells but had no discernible impact on the steady-state heterogeneity or differentiation trajectory of thymocytes. These findings show that VISTA plays a constitutive function in maintaining naïve T cell identity exclusively outside of the thymus. The disruption of T cell quiescence owing to the loss of VISTA was inextricably linked to undermining peripheral T cell tolerance to antigen in polyclonal, transgenic, and endogenous antigen-specific T cell systems. The function of VISTA in vivo could be amplified using anti-VISTA agonists, which augmented T cell tolerance induction in the same systems in part by enhancing peripheral T cell death under costimulation-deficient conditions. Surprisingly, we did not observe a significant impact of VISTA loss or targeting on other modes of T cell suppression, such as anergy induction. A highly important note is that the function of VISTA is relegated to controlling naïve T cell fate because its impact as well as its expression are all but obliterated under inflammatory conditions (e.g., LPS, CFA, and poly-IC) in which CTLA-4, LAG3, and PD-1 play prominent immunoregulatory roles under inflammatory states. However, under tolerogenic conditions, VISTA expression was sustained. Unlike all other coinhibitory molecules expressed after T cell activation, VISTA presents the first of a class of NCRs critical for maintaining naïve T cell quiescence, directing naïve T cell responses to antigen, and peripheral T cell tolerance. In addition, VISTA represents a specific NCR that can be targeted by both agonists and antagonists to impart opposing outcomes on T cell fate.

These insights explain the impact of VISTA loss on exacerbating T cell–directed immune aggression in multiple mouse models such as SLE, GVHD, and experimental autoimmune encephalomyelitis. The gene signature of VISTA loss was predictive in multiple human autoimmune diseases (e.g., lupus and RA), suggesting the therapeutic potential of VISTA agonistic targeting. Our work also presents a high-resolution profile of the earliest stages of thymocyte and T cell differentiation and the landscape of T cell responses under tolerizing versus immunizing settings with antigen.

One of the important remaining questions is the identity of the regulatory networks that constitutively maintain the expression of VISTA in naïve T cells and distinguish it from other established NCRs. Previous work demonstrated that the TFs p53 and HIF1α bind the VISTA promoter and up-regulate VISTA expression (*14*, *72*). Analysis of ImmGen datasets of mature CD4$^+$ T cells shows a direct correlation between VISTA and p53 (fig. S13A). Analysis of the Encyclopedia of DNA Elements (ENCODE) database (*73*) for TF binding revealed multiple potential TFs to the VISTA promoter. We screened the TFs expressed in T cells out of these putative regulators and found that *Fos*, *JunD*, and *NFκB* all have binding sites in the VISTA gene (fig. S13B and table S14). Because VISTA expression is reduced on T cells responding to antigen under conditions of inflammation (Fig. 3), we examined the expression of these potential regulators under tolerization versus inflammatory (immunizing) conditions using the same RNA-seq dataset used in Fig. 3. *Jund* and *Fos* were significantly up-regulated under inflammatory conditions, thereby showing an inverse relationship with VISTA expression (fig. S13C). Indeed, this finding was supported by analysis of independent datasets in the ImmGen database (fig. S13, D and E). This would imply that they are potential TF repressors of VISTA expression. More extensive studies will determine the regulatory networks that distinguish inhibitory checkpoint expression and activity.

Materials and methods

Mice and cell lines

Eight- to 10-week-old C57BL/6 mice WT were purchased from Charles River (Wilmington, MA). B6N.129S5(B6)-*Vsir*^{tm1Lex}/Mmucd (VISTA KO) mice were obtained from the Mutant Mouse Resource & Research Centers (www.mmrrc.org; stock no. 031656-UCD) and were fully backcrossed onto the C57BL/6 and BALB/c backgrounds. VISTA$^{fl/fl}$ mice were bred and screened of VISTA$^{fl/fl}$ mice as described previously (*14*). Conditional deletion of VISTA in the CD4$^+$ T cell compartment was achieved by crossing VISTA$^{fl/fl}$ mice to hemizygous B6.Cg-Tg (Cd4-cre)1Cwi/BfluJ mice (stock no. 022071). Cre-positive mice were compared with Cre-negative littermate controls. Deletion of VISTA on CD4$^+$ T cells and thymocytes was further confirmed by flow cytometry. *Rag1*$^{-/-}$ (B6.129S7-*Rag1*^{tm1Mom}/J) gender-matched 6- to 8-week-old mice were purchased from Charles River. Inducible-deletion of VISTA was achieved by crossing CD4-Cre ER^{T2} (Jackson Laboratory) to VISTA$^{fl/fl}$. Human VISTA KI mice were generated by GenOway (Lyon, France) by knocking in a single copy of human *VISTA* cDNA (GenBank accession no. NM_022153.2) as an in-frame fusion with the 3′ end of the murine signal peptide coding sequence located in the exon 3. This approach resulted in the deletion of part of the exon 3 coding sequence and led to the disruption of the murine gene (fig. S7A). Homologous recombination was done in the C57BL/6-derived ES line. Mouse chimeras were then bred with C57BL/6 Cre deleter mice [Jackson, B6.C-Tg (CMV-cre)1Cgn/J] to excise the neomycin selection cassette (Neo) and to generate heterozygous mice carrying the Neo-excised humanized KI allele. Subsequently, mice were bred to generate homozygous human VISTA KI mice. hVISTA expression and mVISTA deletion were validated by PCR and flow cytometry using fluorophore-conjugated anti-hVISTA [clone 803 (represented in Fig. 2 and fig. S7)] and anti-mVISTA (clone 13F3). *Rag2*$^{-/-}$ OT-II were bred onto B6-Ly5.1/Cr (B6.SJL-*Ptprc*^{a}*Pepc*^{b}/BoyCrCrl, Charles River) for detection using congenic marker. Gender-matched littermates were then used in the adoptive transfer experiments. For specific experiments, hVISTA KI homozygous mice were interbred with *Rag2*$^{-/-}$ OT-II B6-Ly5.1. Act-mOva mice were purchased from Jackson Laboratory (C57BL/6-Tg(CAG-OVA)916Jen/J, stock no. 005145) and were gender-matched with the donor mice. KLF2-GFP mice were developed, bred, and screened at the University of Minnesota (Stephen Jameson lab) (*17*, *18*). TEa transgenic mice [B6.Cg-Tg(Tcra,Tcrb)3Ayr/J] expressing GFP were bred and screened in-house (*74*). CB6F1/J (C57BL/6 × Balb/c) (100007), NZBWF1/J (100008), B6.MRL-Fas^{lpr} (000482), *Bcl2l11*$^{tm1.1Ast}$ Bim-deficient (004525) mice were all purchased from Jackson Laboratory. K/BxN transgenic mice were bred and screened in-house. 2D2 TCR transgenic mice were purchased from Jackson Laboratory and bred onto VISTA$^{-/-}$ B6 background (*13*). Mice were maintained under specific-pathogen–free conditions in the Dartmouth Center for Comparative Medicine and Research. The Animal Care and Use Committee of Dartmouth College approved all animal experiments. For experiments

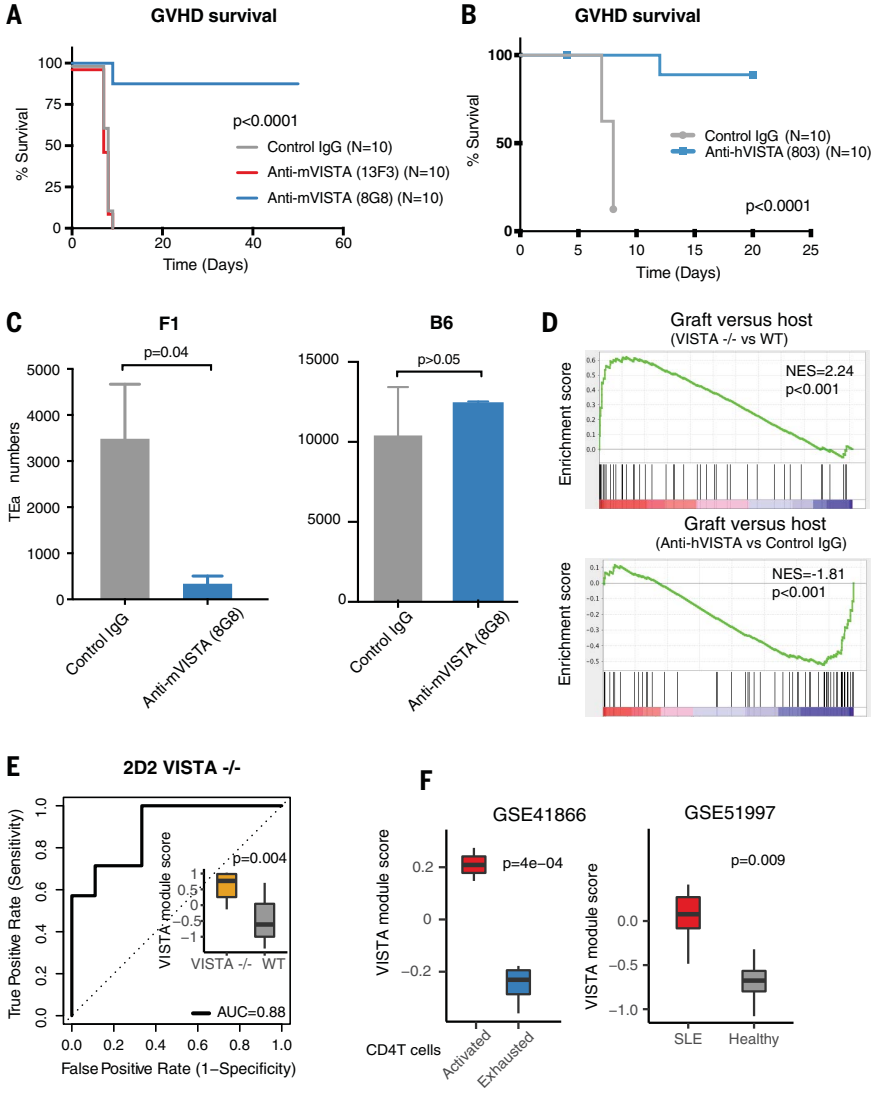


Fig. 4. VISTA targeting induces systemic tolerance and T cell deletion, whereas VISTA deletion imparts an autoimmune-associated gene signature. (**A**) C57BL/6 bone marrow (BM) and splenocytes (10^7 each) were transferred to lethally irradiated BALB/c recipient mice and treated with antagonist (clone 13F3, red) or agonist (clone 8G8, blue) anti-mVISTA antibodies (200 μg per mouse) on day 0 to induce acute GVHD. Survival was monitored for the indicated time periods. (**B**) C57BL/6 BM and hVISTA-expressing splenocytes (10^7 each) were transferred to lethally irradiated BALB/c recipient mice, and anti-hVISTA agonist (clone 803, blue) or IgG control (200 μg per mouse) was administered on day 0 to induce GVHD. Survival was monitored. In all survival experiments, *P* values were calculated by log rank test. Data from the GVHD models are representative of two independent experiments with at least 10 mice per group. (**C**) TEa $CD4^+$ T cells (2×10^6 cells per mouse) were transferred into 650-centigray-irradiated F1 hosts (left) or age- and-gender matched C5/B6 hosts (right) and intravenously treated with an anti-mVISTA antibody (clone 8G8) or control IgG followed by analysis of cell numbers on day 2 posttransfer. Data show the mean numbers of recovered TEa T cells for each treatment as percentages of $CD4^+$ T cells. Data are representative of four independent experiments with five mice per group. Each bar refers to the mean value, and each error bar refers to one SD; all *P* values were calculated by Student's *t* test. (**D**) GSEA pathway enrichment plot indicating the GVHD gene set enriched in $VISTA^{-/-}$ versus WT (top) and anti-hVISTA (clone 803)–treated versus control IgG–treated mice (bottom, obtained from Fig. 2 data). NESs and *P* values are shown for each gene set. *P* values were calculated by Kolmogorov-Smirnov test. (**E**) Receiver operating characteristic (ROC) curves for predicting VISTA mutation status in 2D2 transgenic $CD4^+$ T cells using VISTA-deficiency module score as the predictor. VISTA-deficiency module defines the gene signature resulting from loss of VISTA on the naïve T cell. The inset is a boxplot depicting the difference in the VISTA module scores for 2D2 transgenic $CD4^+$ T cells in $VISTA^{-/-}$ and WT mice. *P* values were calculated by Wilcoxon rank sum test. (**F**) Boxplots showing the difference in VISTA module scores between exhausted versus activated $CD4^+$ T cells (left) and $CD4^+$ T cells from SLE patients and healthy individuals (right). *P* values were calculated by Wilcoxon rank sum test.

involving thymocytes, gender- and age-matched littermates from 3- to 4-week-old mice were used. Both male and female mice were used in independent experiments. A Jurkat cell line expressing VISTA was generated using the construct pEF1a-hVISTA-IRES-ZsGreen1. The pEF1-IRES-ZsGreen1 construct was initially purchased from TakaraBio (cat. 631976), and hVISTA sequence was cloned. The Jurkat cell line (ATCC, TIB-152, clone E6-1) was then transfected with the construct using cell line Nucleofector (Lonza, VCA-1003) following the manufacturer's protocol. A stable pool was then generated, and hVISTA expression on this cell line compared with control Jurkat cells transfected with empty vector was assessed using fluorophore-conjugated anti-VISTA clone 803.

Antibodies

Antibodies for mouse and human VISTA were generated as described previously (*63*, *75*). Female C57BL/6 mice were immunized with human VISTA–Ig fusion protein emulsified in complete Freund's adjuvant (CFA). They were boosted 4 weeks later with protein in incomplete Freund's adjuvant (IFA) and then 6 weeks later with A20 cells overexpressing VISTA–red fluorescent protein. Finally, they were boosted with VISTA-Ig fusion protein without the adjuvant. Four days after the last boost, spleens from immunized mice were provided to APS Ltd. Hybridomas and antibodies were generated by APS Ltd. Hybridoma clones that produced VISTA-specific antibodies were selected after limiting dilution and screened by both ELISA and flow cytometry methods. Anti-hVISTA clone 803 was humanized into a full IgG2 human antibody by Aragen Biosciences. To demonstrate specificity of the clone anti-hVISTA 803, 10^6 peripheral blood mononuclear cells (PBMCs) were stained with 5 μg/ml of anti-hVISTA 803 in the presence of 10 μg of soluble VISTA-Ig. In addition, the Jurkat cell line was stably transfected with human VISTA, and staining was compared with control vector–transfected Jurkat cells. Primary immune cell subsets from human peripheral blood and multiple mouse tissues were stained with both hVISTA and mVISTA antibody clones to demonstrate specificity (fig. S7). Anti-hVISTA 803 is a chimeric human IgG2 antibody. Both antagonist anti-mVISTA clone 13F3 and agonist anti-mVISTA clone 8G8 are hamster IgG clones, and monoclonal hamster

IgG (BioXCell, Lebanon, NH) was used as their control.

Adoptive cell transfer

For all experiments involving adoptive transfer, naïve $CD4^+$ T cells from donor age- and sex-matched mice were purified using a naïve $CD4^+$ T cell isolation kit (Miltenyi). For experiments involving Ly5.1 WT and CD4-Cre × $VISTA^{fl/fl}$ T cell transfers (Fig. 1 and fig. S5) of naïve $CD4^+$ T cells, the donor cells were mixed at a 1:1 ratio (validated by flow cytometry) and then a total of 1×10^6 cells were adoptively co-transferred by intravenous tail vein injection into recipient *Rag1*$^{-/-}$ hosts. Mice were then injected with 5 μg of either hamster anti-CD3ε or hamster IgG control (BioXcell). Cells were recovered on day 5 posttransfer, and ratios were quantified by flow cytometry using congenic markers. Inducible deletion of VISTA was achieved by i.p. injection of tamoxifen, as recommended by Jackson Laboratory (https://www.jax.org/research-and-faculty/resources/cre-repository/tamoxifen). Briefly, three injections of tamoxifen were required before full deletion of VISTA on the CD4 T cell compartment was observed by flow cytometry. Cells were then isolated as described and adoptively transferred. For experiments involving OT-II and hVISTA OT-II adoptive transfers, single group transfers (3×10^6 cells per mouse) of congenically discordant $CD45.1^+$ OT-II cells into either Act-Ova or B6 were performed and treated with either anti-VISTA (200 μg) or IgG control followed by cell recovery and quantification 48 hours after transfer. TEa transgenic $CD4^+$ T cells (2×10^6) were transferred into cGy 650 irradiated F1 hosts, and mice were either treated with anti-mVISTA (8G8) or hamster IgG control (200 μg/mouse) followed by cell recovery 48 hours after transfer. For TEa quantification in multiple tissues, cells were transferred under the same conditions and isolated from each of the aforementioned tissues. Isolation from spleen and lymph nodes followed the standard procedure (*61*, *76*). Isolation from liver and lung tissues required Percoll density centrifugation, whereas isolating lymphocytes from small intestine followed the described procedure (*77*). Isolation from bone marrow was performed as described (*78*). TEa cell numbers were then quantified by flow cytometry using GFP expression in addition to Thy1.1 (clone OX-7) and CD4 staining.

Acute GVHD model

For the mVISTA treatment experiments, 10-week-old BALB/c recipients and C57BL/6 donor mice were purchased from Charles River. Recipient mice were subjected to total body irradiation (TBI) emanating from a cesium-137 source twice at 450 centigray (cGy) at D0 (9:30 a.m. and 1:30 p.m.) before transfer. Donor mice were euthanized, and bone marrow was harvested by flushing femur and tibia with HBSS. Red blood cells were lysed using ACT solution, and a single-cell suspension of splenocytes and BM cells was prepared and counted. Recipient mice received 10 million bone marrow cells and 10 million spleen cells along with 200 ug of control IgG or anti-mVISTA agonist clone 8G8 or antagonist clone 13F3. Cells and antibodies were administered by tail vein intravenous injection. Mice were weighed regularly to monitor disease progression. Mice were euthanized when they showed signs of morbidity. For anti-hVISTA experiment, the same procedure was applied with the exception that hVISTA splenocytes and WT BM cells were intravenously injected and the mice were either treated with anti-hVISTA 803 or IgG2 control.

Antigen tolerization and immunization

As described previously (*58*, *62*), intravenous injection of soluble 2w1s:I-A^b (EAWGALANWAVDSA) antigen was used to induce antigen-specific T cell tolerance, whereas injection of antigen in the presence of LPS adjuvant was used to provide an immunizing inflammatory condition. To induce T cell tolerance, two doses of 100 μg of 2w1s peptide (Genscript Corp) were intravenously injected on days 0 and 3, followed by analysis on day 7. For MOG antigen tolerization, MOG_{35-55} peptide (200 μg) was intravenously injected on day 0, followed by analysis on day 7. For immunization, mice were intravenously injected on day 0 with 2w1s (100 μg) and LPS (5 μg), followed by analysis on day 7. In the 2w1s peptide tolerization scRNA-seq (Figs. 2 and 3), cells were analyzed 72 hours post intravenous injection. Anti-VISTA or IgG control treatments (200 μg per mouse) were injected on day 0.

Tetramer enrichment

Staining with tetramer and enrichment for antigen-specific endogenous T cell quantification were performed as described previously (*61*, *62*, *76*). Spleen and lymph nodes (inguinal, axillary, brachial, cervical, mesenteric, and periaortic) were harvested for each mouse. A single-cell suspension was prepared in 200 μl of Fc-block supplemented sorter buffer (Fc block + 2% BSA, 0.05% sodium azide). PE-conjugated 2w1s or MOG_{35-55} tetramers (MBL international) was added at a concentration of 20 nM, and the cells were incubated for 1 hour at RT, followed by washing with 15 ml of ice-cold sorter buffer (PBS + 2% BSA, 0.1% sodium azide). The tetramer-stained cells were then resuspended in a volume of 200 μl of sorter buffer, mixed with 50 μl of anti-PE antibody conjugated magnetic microbeads (Miltenyi Biotec), and incubated on ice for 20 min, followed by two washes with 10 ml of sorter buffer. The cells were then resuspended in 3 ml of sorter buffer and passed over a magnetized LS column (Miltenyi Biotech). The column was washed with 3 ml of sorter buffer three times and then removed from the magnetic field. The bound cells were eluted by pushing 5 ml of sorter buffer through the column with a plunger. The resulting enriched fractions were resuspended in 0.1 ml of sorter buffer; a small volume was removed for cell counting, and the rest of the sample was stained with a cocktail of fluorochrome-labeled antibodies specific for B220, CD19 CD11b, CD11c, F4/80, CD3, CD8, NK1.1, CD4, and CD44. Quantification of the number of 2w1s:I-A^b cells per mouse followed the protocol described (*76*).

Flow cytometry and staining

Gentle manual dissociation of splenocytes and lymph node cells to single-cell suspensions was performed as described previously (*61*, *76*). For thymocyte staining for analysis and sorting (fig. S3), thymi were collected in 5 ml of HBSS supplemented with collagenase/DNase I (Worthington) and homogenized gently then incubated for 15 minutes (37°C, 5% CO_2). Cells were then washed and stained for flow cytometric analysis or sorting. For all flow cytometry experiments, T cells were stained with a fixable live-dead stain (Invitrogen) in PBS followed by surface antibody staining in FACS buffer (PBS with 0.5% BSA and 0.1% sodium azide). For intracellular cytokine staining (fig. S9), cells were incubated for 4 hours at 37°C in RPMI-1640 medium plus 10% FBS in the presence of 10 ng/ml PMA (Sigma-Aldrich), 1 μM ionomycin (EMD Chemicals), and 10 μg/ml brefeldin A (Sigma-Aldrich). Cells were then stained for surface markers as described. Intracellular staining was preformed using the eBioscience Cytofix/Cytoperm kit (Thermofisher). Surface-stained cells were then stained with anti–IL-2 (JESS-5H4) and anti–IFN-γ (XMG1.2). The anergic phenotype of 2w1s:I-A^b $CD4^+$ T cells and $Foxp3^+$ T_{reg} quantification were performed as previously described (*65*). Briefly, tetramer enrichment was performed as described above followed by the same surface-staining procedure for tetramer experiments in addition to staining for CD73 (eBioTY/11.8), FR4 (eBio12A5). Stained cells were then treated with eBioscience Foxp3 fixation and permeabilization buffer sets (Thermofisher) following the manufacturer's instructions and stained for Foxp3 (FJK-16s). Anergic 2w1s:I-A^b $CD4^+$ T cells were quantified as tet^+ $CD44^{hi}$ $Foxp3^-$ $CD73^{hi}$ $FR4^{hi}$. $Foxp3^+$ thymocytes were also stained using the same kit in addition to VISTA and CD4 surface staining. For naïve T cell and thymocyte VISTA surface staining, clone MIH-63 (Biolegend) was used. Samples were collected on MACSQuant Analyzer 10 (Miltenyi Biotec) and analyzed using FlowLogic Software 7.2 (Miltenyi Biotec). For experiments using OT-II $CD4^+$ T cells (Fig. 2 and figs. S6 and S7), the donor T cells were stained with Vβ5 (MR9-4) and CD45.1 (A20) in addition

to CD4 staining. Dead cells were calculated as the percentage of total OT-II recovered using viability dye (Near IR, Invitrogen). For analysis of frequencies of different immune populations after anti-VISTA treatment (figs. S6J and S7G), the following antibody clones were used: CD11b (M1/70), Ly6G (1A8), Ly6C (HK1.4), CD3 (17A2), CD4 (RM4-5), CD8 (53-6.7), TCRβ (H57-597), F4/80 (BM8), CD19 (6D5), NK1.1 (PK136), CD11c (N418), and Siglec H (551). Neutrophils were identified as $CD11b^+$ $Ly6G^+$ $Ly6C^-$, whereas monocytes were identified as $CD11b^+$ $Ly6C^+$ $Ly6G^-$. Macrophages were gated as $CD11b^+$ $F4/80^+$ cells. $CD4^+$ and $CD8^+$ T cells were pregated on $CD3^+$ $TCR\beta^+$ live cells. For dendritic cells (DCs), spleens were digested and processed as described previously (*79*) and a lineage gating was added (CD19 CD3 Ly6G NK1.1). Conventional DCs were defined as $Lineage^{-ve}$ $CD11c^+$ live cells, whereas plasmacytoid DCs (pDCs) were defined as $Lineage^{-ve}$ $CD11c^{int}$ Siglec H^+ live cells.

Flow sorting for single-cell sequencing and total RNA-seq

For scRNA-seq experiments depicted in Fig. 1A and scATAC-seq in Fig. 1D: Cells were stained with CD4 (clone RM4-5), CD62L (MEL-14) and CD44 (IM-7), and lineage/Dump (Lin) gate (CD11b, CD11c, NK1.1, CD19, F4/80, CD8, B220) for 20 min on ice, washed, and then flow-sorted using FACS-ARIA II (BD Biosciences) for $CD4^+$ $CD44^{lo}$ (lowest 20% $CD44^-$) $CD62L^{hi}$ Lin^- cells into 96-well plates. This same procedure was applied for fig. S5 except that cells were sorted based on $CD4^+$ $CD44^{hi}$ Lin^{-ve}. For thymocyte sorting in fig. S3, cells were stained with CD4, CD8 (clone: 53-6.7), and lineage (CD11b, CD11c, NK1.1, CD19, F4/80). For scRNA-seq of 2w1s:$I\text{-}A^b$, cells were first stained and enriched with tetramer, then flow sorted using the staining panel (CD4, CD44, CD8, 2w1s-Tet) for $CD4^+$ $CD44^{hi}$ Tet^+ $CD8^-$ Lin^- and lineage was defined as (CD11c, B220, CD19, CD11b, F4/80, NK1.1). For experiments depicted in fig. S3, A to C, $CD4^+$ $CD44^{lo}$ $CD62L^{hi}$ from KLF2-GFP mice were sorted based on GFP reporter expression with $KLF2^{hi}$ defined by 20% highest expression and the $KLF2^{lo}$ defined as 20% lowest (positive) expression. For sorting of $VISTA^{hi}$ versus $VISTA^{lo}$, cells were stained according to the same procedure, in addition to VISTA (clone MIH-63). Purity was validated by using subsequent flow cytometry and scRNA-seq analysis in which nonspecific cells were excluded.

Single-cell RNA sequencing and normalization

Droplet-based 5′-end scRNA-seq was performed by the 10x Genomics platform, and libraries were prepared by the Chromium Single Cell 5′ Reagent kit according to the manufacturer's protocol (10x Genomics, CA, USA). The Cell Ranger Single-Cell Software Suite (10x Genomics) was used to perform barcode processing and transcript counting after alignment to the mm10 reference genome with default parameters. The Seurat R package (*80*) was applied to filter out low-quality cells, normalize gene expression profiles, and cluster cells. Cells expressing >10% mitochondrial gene counts or expressing less than 500 genes were discarded using the *FilterCells* function. Then, the *NormalizeData* function was applied to normalize and log transform the raw counts for each cell on the basis of its library size.

Single-cell unsupervised clustering

The normalized expression matrices of naïve $CD4^+$ T cells, $CD4^+$ $CD44^{hi}$ MP T cells, $CD4^+$ thymocytes, and 2w1s:$I\text{-}A^b$ specific $CD4^+$ T cells were processed by filtering the nonexpressed genes separately. The unsupervised clustering was applied in each dataset as follows: (i) Top variant genes with dispersion higher than 0.5 and average expression higher than 0.15 were selected and used as the input for principal components analysis (PCA) to reflect the major biological variation in the data. (ii) The top 15 PCs were chosen for t-SNE dimension reduction by the *RunTSNE* function and unsupervised clustering. Specifically, the *FindClusters* function was used to cluster the cells. (iii) After the cell clusters were determined, marker genes for each cluster were identified by the *FindAllMarkers* function with the default parameter. The biological annotation of each cluster was further described by the marker gene function reported in the literature and the pathways specifically associated with the cluster (see "Pathway enrichment analysis") or the representation of the marker gene expression in the ImmGen database, which has a clear description of different $CD4^+$ T subsets (*81*). We examined the expression pattern of Z-transformed average gene expression of cluster marker genes in ImmGen $CD4^+$ T cells. The ImmGen $CD4^+$ T cell lineage with highest expression level of the cluster marker genes was chosen as the annotation for the $CD4^+$T cell cluster.

Identification of autoreactive T cells

Droplet-based 5′-end single-cell TCR sequencing (scTCR-seq) was performed by the 10x Genomics platform, and libraries were prepared by the Chromium Single Cell Immune Profiling Solution kit according to the manufacturer's protocol (10x Genomics, CA, USA). The Cell Ranger Single-Cell Software Suite VDJ pipeline (10x Genomics) was used to perform barcode processing and consensus TCR annotation after alignment to the mm10 reference genome with default parameters. The annotated TCR sequences of naïve $CD4^+$ T cells and $CD4^+$ $CD44^{hi}$ MP T cells from $VISTA^{-/-}$ and WT mice and of $CD4^+$ T cells from Fas lpr and Bim-deficient mice were processed by filtering out nonproductive TCRs. To identify the autoreactive $CD4^+$ T cells, the Vβ CDR3 sequences in naïve $CD4^+$ T cells were aligned with the Vβ CDR3 sequences in $CD4^+$ $CD44^{hi}$ MP T cells and $CD4^+$ T cells from Fas lpr and Bim-deficient mice. The *pairwiseAlignment* function in the Biostring R package (*44*) with parameter "type = local, gapOpening = 10, gapExtension = 4" was used for sequence matching.

Developmental trajectory inference

To determine the potential lineage differentiation between $VISTA^{-/-}$ and WT, Monocle (version 2) (*82*) algorithm was used with scRNA thymus double-positive, single-positive, and naïve $CD4^+$ T cells raw counts matrix as the input. The *newCellDataset* function was used to build a CellDataSet object with the parameter "expressionFamily = negbinomial." Then, differential gene expression analysis was performed using the *differentialGeneTest* function with the parameter "fulModelFormulStr = ~Cluster_assign, reducedModelFomulaStr = ~batch." Specifically, "Cluster_assign" refers to the cluster identification of the scRNA and batch refers to the batch experiment number during which the scRNA was sequenced. Moreover, the dimension reduction was performed by the *reduceDimension* function with the parameter "max_components = 2, reducedModelFormulaStr = ~ batch, method = DDRTree." The differentiation trajectory was then inferred with the default setting of Monocle.

Pathway enrichment analysis

The differentially expressed genes between different cell clusters or different VISTA perturbations (e.g. $VISTA^{-/-}$ or anti-VISTA treatments) were ranked on the basis of the average log-fold change. To annotate the pathways that were involved in the differentially expressed genes, pathway gene sets were downloaded from the C2 category of the Molecular Signatures Database (MSigDB v6.2) database (*83*). Furthermore, gene sets with less than 10 effective genes (i.e., the number of genes presented in a gene expression dataset) were discarded. The preranked gene set enrichment analysis (GSEA) software was used to calculate the enrichment of each pathway in the genes that are most informative in each gene list.

2D2 $CD4^+$ T cell isolation for total RNA-seq

Naïve $CD4^+$ T cells were isolated from 4- to 8-week-old asymptomatic male and female 2D2 transgenic mice on the $VISTA^{-/-}$ or WT background (*13*) using the naïve CD4 T cell isolation kit (Miltenyi Biotec). Cell purity (~96 to 98%) and viability were assessed by flow cytometry.

Calculation of VISTA module score

Differential gene expression analysis was performed for each gene between grouped CD4-Cre $VISTA^{-/-}$ and WT naïve $CD4^+$ T cells using

Wilcoxon rank sum test. A P value of <0.05 was used as the threshold to determine the statistical significance, and the log-fold change was used to determine if the gene was up- or down-regulated in the VISTA$^{-/-}$ naïve CD4^{+} T cells (table S13). The VISTA gene module was defined as the combination of significantly up- and down-regulated genes in VISTA$^{-/-}$ naïve CD4^{+} T cells. In a given gene expression dataset, the VISTA module score was first calculated as the average gene expression difference of up- and down-regulated genes in the module and then Z-transformed into normal distribution. A higher VISTA module score indicated a higher chance of VISTA deficiency in a given CD4^{+} T cell. Validation of the VISTA module score was performed in an independent 2D2 transgenic CD4^{+} T cells RNA-seq dataset. Area under the ROC (AUC) was used as a metric for evaluating the accuracy of the VISTA module score in capturing VISTA deficiency. For each CD4^{+} T cell sample, a threshold was set beginning with the lowest score; all samples with a score higher the threshold were predicted to be VISTA$^{-/-}$, and all samples below the threshold were predicted to be WT. The sensitivity and specificity were then calculated for each threshold by comparing the predicted VISTA$^{-/-}$ with the actual VISTA$^{-/-}$.

RNA-seq alignment for KLF2^{hi} versus KLF2^{lo}, VISTA^{hi} versus VISTA^{lo} naïve CD4^{+} T cells, and 2D2 VISTA$^{-/-}$ CD4^{+} T cells total RNA-seq

Sequencing was performed on a NextSeq 500 (Illumina) instrument to obtain an average of raw 100-bp single end reads per sample. Raw .bcl files were demultiplexed using the Illumina bcl2fastq2 pipeline. The quality of the fastq files was examined with the *FastQC* software (www.bioinformatics.babraham.ac.uk/projects/fastqc). Raw fastq files were trimmed using the software *Trimmomatic* by setting the parameter "SLIDINGWINDOW: 4:15 LEADING: 3 TRAILING: 3 MINLEN: 36." The trimmed fastq files were than aligned to the mouse mm10 reference genome and normalized to obtain transcripts per kilobase million (TPM) for each RNA-seq sample using the software *Salmon* with the parameter "-l A" (*84*).

Single-cell ATAC sequencing and normalization

CellRanger-atac v1.1 was used to generate fastq files (mkfast) and to demultiplex, align to the mouse mm10 genome, and call peaks using the "count" pipeline (http://software.10xgenomics.com/single-cell/overview/welcome). Peak count matrices were aggregated using the "aggr" function and normalized to sequencing depth. Cells with peak counts higher than 5000 were kept for further analyses. To further examine the quality of the scATAC-seq, The fragment file, which records the full list of all unique fragments across all cells, was used for quality control. Specifically, the fraction of fragments in total peaks was calculated by the number of fragments that mapped to the peak region divided by the total number fragments in each cell. The blacklist fragments ratio was calculated by the number of fragments that mapped to the blacklist region versus the number of fragments that mapped to peak region (*85*). As recommended by Stuart *et al.* (*86*), cells having total number of fragments in peaks higher than 1000 and fraction of peaks that located in the peak higher than 15% and blacklist ratio lower than 2.5% were considered as good cells. 99.36% of cells passed the quality control.

The generated peak matrix was binarized, and then we performed the term frequency–inverse document frequency ("TF-IDF") transformation as suggested by Cusanovich *et al.* (*87*). We first divided each peak in each cell by the total number of accessibility of peaks in the cell (the "term frequency") and then multiplied these values by the inverse accessibility of the peaks across cells (the "inverse document frequency").

Single-cell ATAC unsupervised clustering

Peaks having at least 100 reads across cells were considered variable peaks for unsupervised clustering. The TF-IDF matrix was used as input to conduct SVD to return LSI components. These steps were performed in the *RunLSI* function in Seurat. We retained 50 dimensions and created a new Seurat object. The clusters were identified using Seurat's SNN graph clustering using the *FindClusters* function and visualized using the *RunUMAP* function (*86*). To identify cluster markers, the binarized peak matrix was used as input to create a CellDataSet object through the *newCellDataset* function with the parameter "expressionFamily = binomialff." Then, the *differentialGeneTest* function with the parameter "fulModelFormulStr = ~Cluster_assign" was used to identify the marker peaks for each cluster. "Cluster_assign" refers to the cluster identification of the scATAC seq (*87*). To further confirm the statistical significance of marker peaks, the *FindMarkers* function in Seurat was used to perform the likelihood ratio test with the parameter "test.use ="LR", laten.vars="peak_region_fragments." The marker peaks function was annotated on the basis of its nearest gene function. The biological annotation of each cluster was further described by the markers peak associated gene function reported in the literature and the calculated gene activity associated with the cluster (see "Gene activity calculation"). The Signac package was used for peak profile visualization [(*86*); https://satijalab.org/signac/]. Specifically, the *CoveragePlot* function grouped the peaks in each cluster and normalized the peaks by sequencing depth and number of cells in each cluster for visualization.

Gene activity calculation

The Cicero package was used to calculate gene activity scores (GA scores) as previously described (*88*). The binary filtered peak counts matrix was used to build a CellDataSet object with the parameter "expressionFamily = binominalff()." Then a cicero_cds object was made using the function *make_cicero_cds* with the parameter "reduced_coordinates = UMAP_coords." Specifically, the UMAP_coords was generated by previous dimension reduction step. The *run_cicero* function was used to calculate the coaccessed peak-to-peak links across all cells with default parameters in the mouse mm10 genome. The *build_gene_activity_matrix* and *normalize_gene_activities* were used to calculate and normalize the gene activity scores for each cell. For visualization convenience, the GA score was transformed into log10(GA score*1000 + 1).

Identification of VISTA expression regulators

All ENCODE transcription factor ChIP-seq bigWig files were accessed and downloaded from the ENCODE official website (https://www.encodeproject.org) (*89*). With a threshold of P value 0.05, the TIP probabilistic method (*90*) was used to determine the potential transcription factors that bind *VISTA* in each cell line.

ELISA for anti-mVISTA isotype determination

Anti-VISTA antibodies were diluted in PBS and coated overnight on an ELISA plate (RND). The blocking step was performed using PBS (1% BSA). This was followed by incubation with anti-hamster clones IgG1, IgG2, IgG2/3, and IgG2/3/4 (BD Biosciences). Anti-mouse IgG1-HRP followed by TMB substrate solution was used for detection.

Assessment of anti-VISTA agonist suppressive properties

NZBW/F1 lupus

Twenty-four–week-old NZB/W F1 mice were treated three times a week with either anti-mVISTA 8G8 or control IgG (200 μg). Proteinuria levels (mg/dl) were recorded weekly using Chemstrip test trips (Roche Diagnostics).

ConA acute hepatitis

Con A (Sigma-Aldrich) was dissolved in PBS and administered in a total volume of (15 mg/kg) 300 μl by intravenous tail vein injection. Mice received anti-VISTA 8G8 or control IgG (200 μg) through intraperitoneal injections 3 hours before Con A injection. Mice were monitored for survival.

K/BxN arthritis

Mice were injected with 100 μl of K/BxN serum on days 0 and 2. Anti-VISTA 8G8 or control

IgG (200 μg) were given every 3 days starting day 0. Clinical scoring was done as previously described (*91*).

Imiquimod induced psoriasis

Fifty mg of 3.5% IMQ cream was prepared by diluting the 5% IMQ cream (Taro Pharmaceuticals, New York, NY) using the vehicle cream (Vanicream; Pharmaceutical Specialties, Cleveland, GA). Imiquimod was applied to the ear of mice daily. At day 14, anti-mVISTA 8G8 or control IgG (200 μg) were administered intraperitoneally every other day. Ear thickness was measured by using an Ozaki caliper (model G-A1-0.4N) (Neill-Lavielle Supply, Louisville, KY).

Impact of anti-VISTA on uptake of apoptotic cells

The procedure followed the following references (*92*, *93*). Thymocytes (single-cell suspension of 10^7 cells/ml) were isolated from C57BL/6 WT mice and exposed to dexamethasone under cell culture conditions (0.1 μM) in complete DMEM for 24 hours to induce apoptosis. This procedure allowed for ~100% apoptosis, which was assessed by annexin V/Propidium idodide (PI) staining followed by flow cytometric analysis. After apoptosis induction, thymocytes were washed twice, then resuspended in pH rodo dye (final concentration of 20 ng/ml) and incubated for 30 min at RT. This is a pH-sensitive dye that emits red dye only in lower pH, such as that located in a phagosome (pH ~5), so it was used to distinguish engulfed thymocytes from unengulfed controls. Thymocytes (1×10^6 cells) were incubated with thioglycolate-elicited peritoneal macrophages (2×10^5 cells) for 90 min in the presence of anti-VISTA (8G8) or control IgG (10 μg/ml). Phagocytosis was determined using flow cytometry by measuring the positive pHrodo-containing ($CD11b^+$ $F4/80^+$) macrophages. $VISTA^{-/-}$ peritoneal macrophages were used as control on the basis of previous work (*14*).

REFERENCES AND NOTES

1. L. Chen, D. B. Flies, Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nat. Rev. Immunol.* **13**, 227–242 (2013). doi: 10.1038/nri3405; pmid: 23470321
2. D. M. Pardoll, The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* **12**, 252–264 (2012). doi: 10.1038/nrc3239; pmid: 22437870
3. E. Ahn *et al*., Role of PD-1 during effector CD8 T cell differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4749–4754 (2018). doi: 10.1073/pnas.1718217115; pmid: 29654146
4. P. J. Noel, L. H. Boise, C. B. Thompson, Regulation of T cell activation by CD28 and CTLA4. *Adv. Exp. Med. Biol.* **406**, 209–217 (1996). doi: 10.1007/978-1-4899-0274-0_22; pmid: 8910687
5. C. T. Kuo, M. L. Veselits, J. M. Leiden, LKLF: A transcriptional regulator of single-positive T cell quiescence and survival. *Science* **277**, 1986–1990 (1997). doi: 10.1126/science.277.5334.1986; pmid: 9302292
6. D. Tzachanis *et al*., Tob is a negative regulator of activation that is expressed in anergic and quiescent T cells. *Nat. Immunol.* **2**, 1174–1182 (2001). doi: 10.1038/ni730; pmid: 11694881
7. I. Yusuf, D. A. Fruman, Regulation of quiescence in lymphocytes. *Trends Immunol.* **24**, 380–386 (2003). doi: 10.1016/S1471-4906(03)00141-8; pmid: 12860529
8. M. Berger *et al*., An *Slfn2* mutation causes lymphoid and myeloid immunodeficiency due to loss of immune cell quiescence. *Nat. Immunol.* **11**, 335–343 (2010). doi: 10.1038/ni.1847; pmid: 20190759
9. A. F. Buckley, C. T. Kuo, J. M. Leiden, Transcription factor LKLF is sufficient to program T cell quiescence via a c-Myc-dependent pathway. *Nat. Immunol.* **2**, 698–704 (2001). doi: 10.1038/90633; pmid: 11477405
10. L. Gorelik, R. A. Flavell, Abrogation of TGFβ signaling in T cells leads to spontaneous T cell differentiation and autoimmune disease. *Immunity* **12**, 171–181 (2000). doi: 10.1016/S1074-7613(00)80170-3; pmid: 10714683
11. S. Ceeraz *et al*., VISTA deficiency accelerates the development of fatal murine lupus nephritis. *Arthritis Rheumatol.* **69**, 814–825 (2017). doi: 10.1002/art.40020; pmid: 27992697
12. D. B. Flies *et al*., Coinhibitory receptor PD-1H preferentially suppresses $CD4^+$ T cell-mediated immunity. *J. Clin. Invest.* **124**, 1966–1975 (2014). doi: 10.1172/JCI74589; pmid: 24743150
13. L. Wang *et al*., Disruption of the immune-checkpoint VISTA gene imparts a proinflammatory phenotype with predisposition to the development of autoimmunity. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 14846–14851 (2014). doi: 10.1073/pnas.1407447111; pmid: 25267631
14. K. W. Yoon *et al*., Control of signaling-mediated clearance of apoptotic cells by the tumor suppressor p53. *Science* **349**, 1261669 (2015). doi: 10.1126/science.1261669; pmid: 26228159
15. D. B. Flies, T. Higuchi, L. Chen, Mechanistic assessment of PD-1H coinhibitory receptor-induced T cell tolerance to allogeneic antigens. *J. Immunol.* **194**, 5294–5304 (2015). doi: 10.4049/jimmunol.1402648; pmid: 25917101
16. J. Wu, J. B. Lingrel, KLF2 inhibits Jurkat T leukemia cell growth via upregulation of cyclin-dependent kinase inhibitor p21WAF1/CIP1. *Oncogene* **23**, 8088–8096 (2004). doi: 10.1038/sj.onc.1207996; pmid: 15361832
17. C. M. Carlson *et al*., Kruppel-like factor 2 regulates thymocyte and T-cell migration. *Nature* **442**, 299–302 (2006). doi: 10.1038/nature04882; pmid: 16855590
18. M. A. Weinreich *et al*., KLF2 transcription-factor deficiency in T cells results in unrestrained cytokine production and upregulation of bystander chemokine receptors. *Immunity* **31**, 122–130 (2009). doi: 10.1016/j.immuni.2009.05.011; pmid: 19592277
19. S. Benzeno *et al*., Cyclin-dependent kinase inhibition by the KLF6 tumor suppressor protein through interaction with cyclin D1. *Cancer Res.* **64**, 3885–3891 (2004). doi: 10.1158/0008-5472.CAN-03-2818; pmid: 15172998
20. G. Narla *et al*., In vivo regulation of p21 by the *Kruppel-like factor 6* tumor-suppressor gene in mouse liver and human hepatocellular carcinoma. *Oncogene* **26**, 4428–4434 (2007). doi: 10.1038/sj.onc.1210223; pmid: 17297474
21. S. E. Baranzini, The role of antiproliferative gene *Tob1* in the immune system. *Clin. Exp. Neuroimmunol.* **5**, 132–136 (2014). doi: 10.1111/cen3.12125; pmid: 25071870
22. M. H. Corjay, M. A. Kearney, D. A. Munzer, S. M. Diamond, J. K. Stoltenborg, Antiproliferative gene BTG1 is highly expressed in apoptotic cells in macrophage-rich areas of advanced lesions in Watanabe heritable hyperlipidemic rabbit and human. *Lab. Invest.* **78**, 847–858 (1998). pmid: 9690562
23. J. P. Rouault *et al*., BTG1, a member of a new family of antiproliferative genes. *EMBO J.* **11**, 1663–1670 (1992). doi: 10.1002/j.1460-2075.1992.tb05213.x; pmid: 1373383
24. G. T. Hart, K. A. Hogquist, S. C. Jameson, Krüppel-like factors in lymphocyte biology. *J. Immunol.* **188**, 521–526 (2012). doi: 10.4049/jimmunol.1101530; pmid: 22223851
25. J. M. Curtsinger, P. Agarwal, D. C. Lins, M. F. Mescher, Autocrine IFN-γ promotes naive CD8 T cell differentiation and synergizes with IFN-α to stimulate strong function. *J. Immunol.* **189**, 659–668 (2012). doi: 10.4049/jimmunol.1102727; pmid: 22706089
26. G. Eisenberg *et al*., Soluble SLAMF6 receptor induces strong $CD8^+$ T-cell effector function and improves anti-melanoma activity in vivo. *Cancer Immunol. Res.* **6**, 127–138 (2018). doi: 10.1158/2326-6066.CIR-17-0383; pmid: 29305520
27. B. C. Miller *et al*., Subsets of exhausted $CD8^+$ T cells differentially mediate tumor control and respond to checkpoint blockade. *Nat. Immunol.* **20**, 326–336 (2019). doi: 10.1038/s41590-019-0312-6; pmid: 30778252
28. S. Kumari *et al*., Cytoskeletal tension actively sustains the T cell immunological synapse. bioRxiv 437236 [Preprint]. 8 October 2018. doi: 10.1101/437236
29. N. Tsopoulidis *et al*., T cell receptor-triggered nuclear actin network formation drives $CD4^+$ T cell effector functions. *Sci. Immunol.* **4**, eaav1987 (2019). doi: 10.1126/sciimmunol.aav1987; pmid: 30610013
30. I. Le Mercier *et al*., VISTA regulates the development of protective antitumor immunity. *Cancer Res.* **74**, 1933–1944 (2014). doi: 10.1158/0008-5472.CAN-13-1506; pmid: 24691994
31. Q. Wang, J. He, D. B. Flies, L. Luo, L. Chen, Programmed death one homolog maintains the pool size of regulatory T cells by promoting their differentiation and stability. *Sci. Rep.* **7**, 6086 (2017). doi: 10.1038/s41598-017-06410-w; pmid: 28729608
32. J. D. Buenrostro *et al*., Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486–490 (2015). doi: 10.1038/nature14590; pmid: 26083756
33. J. F. Ashouri, A. Weiss, Endogenous Nur77 is a specific indicator of antigen receptor signaling in human T and B cells. *J. Immunol.* **198**, 657–668 (2017). doi: 10.4049/jimmunol.1601301; pmid: 27940659
34. J. M. Conley, M. P. Gallagher, L. J. Berg, T cells and gene regulation: The switching on and turning up of genes after T cell receptor stimulation in CD8 T cells. *Front. Immunol.* **7**, 76 (2016). doi: 10.3389/fimmu.2016.00076; pmid: 26973653
35. R. S. Akondy *et al*., Origin and differentiation of human memory CD8 T cells after vaccination. *Nature* **552**, 362–367 (2017). doi: 10.1038/nature24633; pmid: 29236685
36. B. Youngblood *et al*., Effector CD8 T cells dedifferentiate into long-lived memory cells. *Nature* **552**, 404–409 (2017). doi: 10.1038/nature25144; pmid: 29236683
37. X. Feng *et al*., Transcription factor Foxp1 exerts essential cell-intrinsic regulation of the quiescence of naive T cells. *Nat. Immunol.* **12**, 544–550 (2011). doi: 10.1038/ni.2034; pmid: 21532575
38. W. Ouyang, M. O. Li, Foxo: In command of T lymphocyte homeostasis and tolerance. *Trends Immunol.* **32**, 26–33 (2011). doi: 10.1016/j.it.2010.10.005; pmid: 21106439
39. W. F. Wong *et al*., Runx1 deficiency in CD4+ T cells causes fatal autoimmune inflammatory lung disease due to spontaneous hyperactivation of cells. *J. Immunol.* **188**, 5408–5420 (2012). doi: 10.4049/jimmunol.1102991; pmid: 22551552
40. G. Zhou *et al*., Identification of systemically expanded activated T cell clones in MRL/lpr and NZB/W F1 lupus model mice. *Clin. Exp. Immunol.* **136**, 448–455 (2004). doi: 10.1111/j.1365-2249.2004.02473.x; pmid: 15147346
41. A. Madi *et al*., T-cell receptor repertoires share a restricted set of public and abundant CDR3 sequences that are associated with self-related immunity. *Genome Res.* **24**, 1603–1612 (2014). doi: 10.1101/gr.170753.113; pmid: 25024161
42. Y. Tikochinski *et al*., A shared TCR CDR3 sequence in NOD mouse autoimmune diabetes. *Int. Immunol.* **11**, 951–956 (1999). doi: 10.1093/intimm/11.6.951; pmid: 10360969
43. K. S. Kim *et al*., Dietary antigens limit mucosal immunity by inducing regulatory T cells in the small intestine. *Science* **351**, 858–863 (2016). doi: 10.1126/science.aac5560; pmid: 26822607
44. H. Pagès, P. Aboyoun, R. Gentleman, S. DebRoy, Biostrings: Efficient manipulation of biological strings. R package version 2.54.0 (Bioconductor, 2019). doi: 10.18129/B9.bioc.Biostrings
45. J. Craft, S. Peng, T. Fujii, M. Okada, S. Fatenejad, Autoreactive T cells in murine lupus: Origins and roles in autoantibody production. *Immunol. Res.* **19**, 245–257 (1999). doi: 10.1007/BF02786492; pmid: 10493178
46. P. Bouillet *et al*., BH3-only Bcl-2 family member Bim is required for apoptosis of autoreactive thymocytes. *Nature* **415**, 922–926 (2002). doi: 10.1038/415922a; pmid: 11859372
47. C. Trapnell *et al*., The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014). doi: 10.1038/nbt.2859; pmid: 24658644
48. H. I. Aksoylar, K. Lampe, M. J. Barnes, D. R. Plas, K. Hoebe, Loss of immunological tolerance in Gimap5-deficient mice is associated with loss of Foxo in $CD4^+$ T cells. *J. Immunol.* **188**, 146–154 (2012). doi: 10.4049/jimmunol.1101206; pmid: 22106000
49. T. Kawabe *et al*., Memory-phenotype $CD4^+$ T cells spontaneously generated under steady-state conditions exert innate T_H1-like effector function. *Sci. Immunol.* **2**, eaam9304 (2017). doi: 10.1126/sciimmunol.aam9304; pmid: 28783663
50. J. Zhu, H. Yamane, W. E. Paul, Differentiation of effector CD4 T cell populations. *Annu. Rev. Immunol.* **28**, 445–489 (2010). doi: 10.1146/annurev-immunol-030409-101212; pmid: 20192806

51. C. U. Blank *et al*., Defining 'T cell exhaustion'. *Nat. Rev. Immunol.* **19**, 665–674 (2019). doi: 10.1038/s41577-019-0221-9; pmid: 31570879
52. X. Tao, S. Constant, P. Jorritsma, K. Bottomly, Strength of TCR signal determines the costimulatory requirements for Th1 and Th2 $CD4^+$ T cell differentiation. *J. Immunol.* **159**, 5956–5963 (1997). pmid: 9550393
53. E. Esplugues *et al*., Control of T_H17 cells occurs in the small intestine. *Nature* **475**, 514–518 (2011). doi: 10.1038/nature10228; pmid: 21765430
54. R. Hirsch, M. Eckhaus, H. Auchincloss Jr., D. H. Sachs, J. A. Bluestone, Effects of in vivo administration of anti-T3 monoclonal antibody on T cell function in mice. I. Immunosuppression of transplantation responses. *J. Immunol.* **140**, 3766–3772 (1988). pmid: 3286764
55. M. A. ElTanbouly, W. Croteau, R. J. Noelle, J. L. Lines, VISTA: A novel immunotherapy target for normalizing innate and adaptive immunity. *Semin. Immunol.* **42**, 101308 (2019). doi: 10.1016/j.smim.2019.101308; pmid: 31604531
56. M. J. Barnden, J. Allison, W. R. Heath, F. R. Carbone, Defective TCR expression in transgenic mice constructed using cDNA-based α- and β-chain genes under the control of heterologous regulatory elements. *Immunol. Cell Biol.* **76**, 34–40 (1998). doi: 10.1046/j.1440-1711.1998.00709.x; pmid: 9553774
57. B. D. Ehst, E. Ingulli, M. K. Jenkins, Development of a novel transgenic mouse for the study of interactions between CD4 and CD8 T cells during graft rejection. *Am. J. Transplant.* **3**, 1355–1362 (2003). doi: 10.1046/j.1600-6135.2003.00246.x; pmid: 14525595
58. E. R. Kearney, K. A. Pape, D. Y. Loh, M. K. Jenkins, Visualization of peptide-specific T cell immunity and peripheral tolerance induction in vivo. *Immunity* **1**, 327–339 (1994). doi: 10.1016/1074-7613(94)90084-1; pmid: 7889419
59. D. L. Mueller, Mechanisms maintaining peripheral tolerance. *Nat. Immunol.* **11**, 21–27 (2010). doi: 10.1038/ni.1817; pmid: 20016506
60. J. D. Altman *et al*., Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94–96 (1996). doi: 10.1126/science.274.5284.94; pmid: 8810254
61. F. P. Legoux *et al*., $CD4^+$ T cell tolerance to tissue-restricted self antigens is mediated by antigen-specific regulatory T cells rather than deletion. *Immunity* **43**, 896–908 (2015). doi: 10.1016/j.immuni.2015.10.011; pmid: 26572061
62. J. J. Moon *et al*., Naive $CD4^+$ T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity* **27**, 203–213 (2007). doi: 10.1016/j.immuni.2007.07.007; pmid: 17707129
63. L. Wang *et al*., VISTA, a novel mouse Ig superfamily ligand that negatively regulates T cell responses. *J. Exp. Med.* **208**, 577–592 (2011). doi: 10.1084/jem.20100619; pmid: 21383057
64. C. G. Fathman, N. B. Lineberry, Molecular mechanisms of $CD4^+$ T-cell anergy. *Nat. Rev. Immunol.* **7**, 599–609 (2007). doi: 10.1038/nri2131; pmid: 17612584
65. L. A. Kalekar *et al*., $CD4^+$ T cell anergy prevents autoimmunity and generates regulatory T cell precursors. *Nat. Immunol.* **17**, 304–314 (2016). doi: 10.1038/ni.3331; pmid: 26829766
66. R. J. Martinez *et al*., Arthritogenic self-reactive $CD4^+$ T cells acquire an $FR4^{hi}CD73^{hi}$ anergic state in the presence of $Foxp3^+$ regulatory T cells. *J. Immunol.* **188**, 170–181 (2012). doi: 10.4049/jimmunol.1101311; pmid: 22124124
67. S. Z. Josefowicz, L. F. Lu, A. Y. Rudensky, Regulatory T cells: Mechanisms of differentiation and function. *Annu. Rev. Immunol.* **30**, 531–564 (2012). doi: 10.1146/annurev.immunol.25.022106.141623; pmid: 22224781
68. X. Liu *et al*., Genome-wide analysis identifies NR4A1 as a key mediator of T cell dysfunction. *Nature* **567**, 525–529 (2019). doi: 10.1038/s41586-019-0979-8; pmid: 30814730
69. N. Chihara *et al*., Induction and transcriptional regulation of the co-inhibitory gene module in T cells. *Nature* **558**, 454–459 (2018). doi: 10.1038/s41586-018-0206-z; pmid: 29899446
70. D. B. Flies, S. Wang, H. Xu, L. Chen, Cutting edge: A monoclonal antibody specific for the programmed death-1 homolog prevents graft-versus-host disease in mouse models. *J. Immunol.* **187**, 1537–1541 (2011). doi: 10.4049/jimmunol.1100660; pmid: 21768399
71. C. E. Grubin, S. Kovats, P. deRoos, A. Y. Rudensky, Deficient positive selection of CD4 T cells in mice displaying altered repertoires of MHC class II-bound self-peptides. *Immunity* **7**, 197–208 (1997). doi: 10.1016/S1074-7613(00)80523-3; pmid: 9285405
72. J. Deng *et al*., Hypoxia-induced VISTA promotes the suppressive function of myeloid-derived suppressor cells in the tumor microenvironment. *Cancer Immunol. Res.* **7**, 1079–1090 (2019). doi: 10.1158/2326-6066.CIR-18-0507; pmid: 31088847
73. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012). doi: 10.1038/nature11247; pmid: 22955616
74. M. Gonzalez *et al*., The balance between donor T cell anergy and suppression versus lethal graft-versus-host disease is determined by host conditioning. *J. Immunol.* **169**, 5581–5589 (2002). doi: 10.4049/jimmunol.169.10.5581; pmid: 12421935
75. J. L. Lines *et al*., VISTA is an immune checkpoint molecule for human T cells. *Cancer Res.* **74**, 1924–1932 (2014). doi: 10.1158/0008-5472.CAN-13-1504; pmid: 24691993
76. F. P. Legoux, J. J. Moon, Peptide:MHC tetramer-based enrichment of epitope-specific T cells. *J. Vis. Exp.* **2012**, e4420 (2012). doi: 10.3791/4420; pmid: 23117190
77. C. J. Couter, N. K. Surana, Isolation and flow cytometric characterization of murine small intestinal lymphocytes. *J. Vis. Exp.* **2016**, e54114 (2016). doi: 10.3791/54114; pmid: 27213538
78. S. R. Amend, K. C. Valkenburg, K. J. Pienta, Murine hind limb long bone dissection and bone marrow isolation. *J. Vis. Exp.* **2016**, e53936 (2016). doi: 10.3791/53936; pmid: 27168390
79. M. B. Dong, M. J. Rahman, K. V. Tarbell, Flow cytometric gating for spleen monocyte and DC subsets: Differences in autoimmune NOD mice and with acute inflammation. *J. Immunol. Methods* **432**, 4–12 (2016). doi: 10.1016/j.jim.2015.08.015; pmid: 26344574
80. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018). doi: 10.1038/nbt.4096; pmid: 29608179
81. T. S. Heng, M. W. Painter; Immunological Genome Project Consortium, The Immunological Genome Project: Networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008). doi: 10.1038/ni1008-1091; pmid: 18800157
82. X. Qiu *et al*., Single-cell mRNA quantification and differential analysis with Census. *Nat. Methods* **14**, 309–315 (2017). doi: 10.1038/nmeth.4150; pmid: 28114287
83. A. Subramanian *et al*., Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005). doi: 10.1073/pnas.0506580102; pmid: 16199517
84. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017). doi: 10.1038/nmeth.4197; pmid: 28263959
85. H. M. Amemiya, A. Kundaje, A. P. Boyle, The ENCODE Blacklist: Identification of problematic regions of the genome. *Sci. Rep.* **9**, 9354 (2019). doi: 10.1038/s41598-019-45839-z; pmid: 31249361
86. T. Stuart *et al*., Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019). doi: 10.1016/j.cell.2019.05.031; pmid: 31178118
87. D. A. Cusanovich *et al*., A single-cell atlas of in vivo mammalian chromatin accessibility. *Cell* **174**, 1309–1324.e18 (2018). doi: 10.1016/j.cell.2018.06.052; pmid: 30078704
88. H. A. Pliner *et al*., Cicero predicts cis-regulatory DNA interactions from single-cell chromatin accessibility data. *Mol. Cell* **71**, 858–871.e8 (2018). doi: 10.1016/j.molcel.2018.06.044; pmid: 30078726
89. C. A. Sloan *et al*., ENCODE data at the ENCODE portal. *Nucleic Acids Res.* **44**, D726–D732 (2016). doi: 10.1093/nar/gkv1160; pmid: 26527727
90. C. Cheng, R. Min, M. Gerstein, TIP: A probabilistic method for identifying transcription factor target genes from ChIP-seq binding profiles. *Bioinformatics* **27**, 3221–3227 (2011). doi: 10.1093/bioinformatics/btr552; pmid: 22039215
91. P. Monach *et al*., The K/BxN mouse model of inflammatory arthritis: Theory and practice. *Methods Mol. Med.* **136**, 269–282 (2007). doi: 10.1007/978-1-59745-402-5_20; pmid: 17983155
92. M. Miksa, H. Komura, R. Wu, K. G. Shah, P. Wang, A novel method to determine the engulfment of apoptotic cells by macrophages using pHrodo succinimidyl ester. *J. Immunol. Methods* **342**, 71–77 (2009). doi: 10.1016/j.jim.2008.11.019; pmid: 19135446
93. J. Suzuki, D. P. Denning, E. Imanishi, H. R. Horvitz, S. Nagata, Xk-related protein 8 and CED-8 promote phosphatidylserine exposure in apoptotic cells. *Science* **341**, 403–406 (2013). doi: 10.1126/science.1236758; pmid: 23845944

ACKNOWLEDGMENTS

We thank E. J. Wherry and C. M. Burns for the careful review of and insightful comments on the manuscript. Flow cytometry and flow sorting experiments were carried out in DartLab (G. Ward), the Immune Monitoring and Flow Cytometry Shared Resource at the Norris Cotton Cancer Center at Dartmouth, with NCI Cancer Center Support Grant 5P30 CA023108-37. RNA-sequencing experiments were carried out at Dartmouth Medical School in the Genomics Shared Resource (by F. Kolling IV), which was established by equipment grants from the NIH and NSF and is supported in part by a Cancer Center Core Grant (P30CA023108) from the National Cancer Institute. **Funding:** Research was supported by NIH grants R01AR070760 (R.J.N.), R01CA214062 (R.J.N.), 1R21CA227996-01A1 (C.C.), RR180061 (C.C.), R01 HL56067 (B.R.B.), R01 HL 11879 (B.R.B.), and R37 AI34495 (B.R.B.) and Cancer Prevention and Research institute of Texas grant RR180061 (C.C.). **Author contributions:** Conceptualization: R.J.N and M.A.E.; Methodology: R.J.N., M.A.E., and Y.Z.; Investigation: M.A.E., Y.Z., E.N., E.S., J.L.L., B.K., I.L., C.C., X.H., S.C.J., K.A.H., C.P., and M.S.M.; Writing – review & editing: R.J.N., M.A.E., D.M., S.C.J., Y.Z., B.R.B., and C.C.; Resources: R.J.N. and C.C.; Supervision: R.J.N. **Competing interests:** R.J.N. is an inventor on patent applications (10035857, 9631018, 9217035, 8501915, 8465740, 8236304, and 8231872) submitted by Dartmouth College, and patent applications (9890215 and 9381244) submitted by Kings College London and Dartmouth College and a co-founder of ImmuNext, a company involved in the development of VISTA-related assets. These applications cover the use of VISTA targeting for modulation of the immune response. **Data and materials availability:** scRNA-seq data were deposited and are available under BioProject accession numbers PRJNA587711, PRJNA587742, PRJNA587790, PRJNA587769, and PRJNA587564. scATAC-seq data were deposited under BioProject accession number PRJNA587562. All antibodies and mice are available under a Material Transfer Agreement by contacting R.J.N.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/eaay0524/suppl/DC1
Figs. S1 to S13
Tables S1 to S15
References

View/request a protocol for this paper from *Bio-protocol*.

15 May 2019; resubmitted 30 July 2019
Accepted 2 December 2019
10.1126/science.aay0524

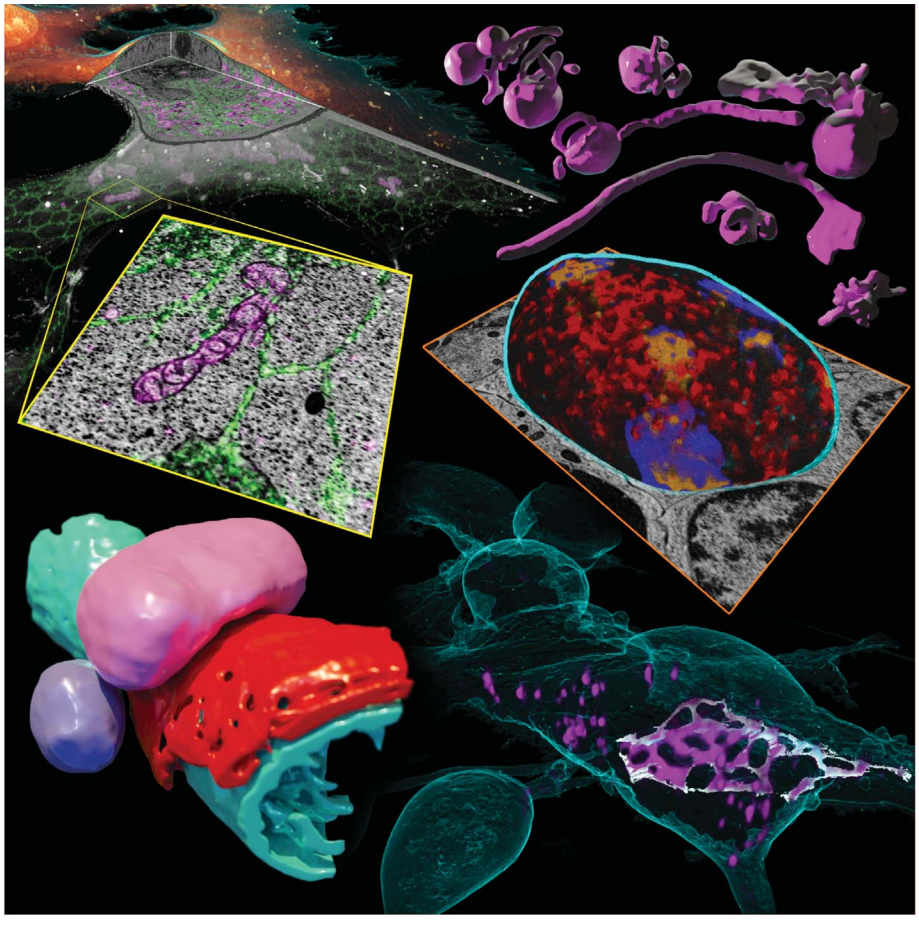
RESEARCH ARTICLE SUMMARY

ADVANCED IMAGING

Correlative three-dimensional super-resolution and block-face electron microscopy of whole vitreously frozen cells

David P. Hoffman*, Gleb Shtengel*, C. Shan Xu, Kirby R. Campbell, Melanie Freeman, Lei Wang, Daniel E. Milkie, H. Amalia Pasolli, Nirmala Iyer, John A. Bogovic, Daniel R. Stabley, Abbas Shirinifard, Song Pang, David Peale, Kathy Schaefer, Wim Pomp, Chi-Lun Chang, Jennifer Lippincott-Schwartz, Tom Kirchhausen, David J. Solecki, Eric Betzig†, Harald F. Hess†

INTRODUCTION: Our textbook understanding of the nanoscale organization of the cell and its relationship to the thousands of proteins that drive cellular metabolism comes largely from a synthesis of biochemistry, molecular biology, and electron microscopy, and is therefore speculative in its details. Correlative super-resolution (SR) fluorescence and electron microscopy (EM) promises to elucidate these details by directly visualizing the nanoscale relationship of specific proteins in the context of the global cellular ultrastructure. However, to date such correlative imaging has involved compromises with respect to ultrastructure preservation and imaging sensitivity, resolution, and/or field of view.



Whole-cell correlative imaging. Cryogenic super-resolution fluorescence microscopy of high-pressure frozen cells coupled with focused ion beam scanning electron microscopy (FIB-SEM) enables multicolor three-dimensional nanoscale visualization of proteins in the context of global ultrastructure. Clockwise from upper left: Volume-rendered cell with correlated orthoslice (inset) of mitochondria and endoplasmic reticulum (ER) proteins; endolysosomal compartments of diverse morphology; heterochomatin subdomains defined by protein reporters of transcriptional activity; adhesion proteins correlated to membrane roughness at contacting cerebellar granule neurons; and a peroxisome (pink) juxtaposed to an ER sheet (red) and mitochondrion (cyan).

RATIONALE: We developed a pipeline to (i) preserve fluorescently labeled, cultured mammalian cells in vitreous ice; (ii) image selected cells in their entirety below 10 K by multicolor three-dimensional structured illumination (3D SIM) and single-molecule localization microscopy (SMLM); (iii) image the same cells by 3D focused ion beam scanning EM (FIB-SEM) at 4- or 8-nm isotropic resolution; and (iv) register all image volumes to nanoscale precision. The pipeline ensures accurate ultrastructure preservation, permits independent optimization of SR and EM imaging modalities, and provides a comprehensive view of how specific subcellular components vary across the cellular volume.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aaz5357

RESULTS: Nearly every system we studied revealed unexpected results: intranuclear vesicles positive for a marker of the endoplasmic reticulum; peroxisomes of increasingly irregular morphology with increasing size; endolysosomal compartments of exceptionally diverse and convoluted morphology; a web-like adhesion network between cerebellar granule neurons; and classically EM-defined domains of heterochromatin and euchromatin each subcharacterized by the presence or absence of markers of transcriptional activity. Two-color cryo-SMLM enabled whole-cell image registration quantifiable down to ~40 nm accuracy. Cryo-SIM, even with its lower resolution, enabled unique discrimination between vesicles of like morphology and aided in segmenting complex 3D structures at FIB-SEM resolution within the crowded intracellular milieu.

CONCLUSION: Our pipeline serves as a powerful hypothesis generator to better understand the findings of biochemistry in the context of the spatially compartmentalized cell. Our approach also carefully preserves the native ultrastructure upon which such hypotheses are based, thus enabling cell-wide or cell-to-cell investigation of the natural variability in protein-ultrastructure relationships. ■

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: hessh@janelia.hhmi.org (H.F.H.); betzige@janelia.hhmi.org (E.B.)
Cite this article as D. P. Hoffman *et al.*, *Science* **367**, eaaz5357 (2020). DOI: 10.1126/science.aaz5357

RESEARCH ARTICLE

ADVANCED IMAGING

Correlative three-dimensional super-resolution and block-face electron microscopy of whole vitreously frozen cells

David P. Hoffman[1]*, Gleb Shtengel[1]*, C. Shan Xu[1], Kirby R. Campbell[2], Melanie Freeman[1], Lei Wang[3,4,5]†, Daniel E. Milkie[1], H. Amalia Pasolli[1]‡, Nirmala Iyer[1], John A. Bogovic[1], Daniel R. Stabley[6], Abbas Shirinifard[7], Song Pang[1], David Peale[1], Kathy Schaefer[1], Wim Pomp[3,4,5]§, Chi-Lun Chang[1], Jennifer Lippincott-Schwartz[1], Tom Kirchhausen[1,3,4,5], David J. Solecki[2], Eric Betzig[1,8,9,10,11,12]#, Harald F. Hess[1]#

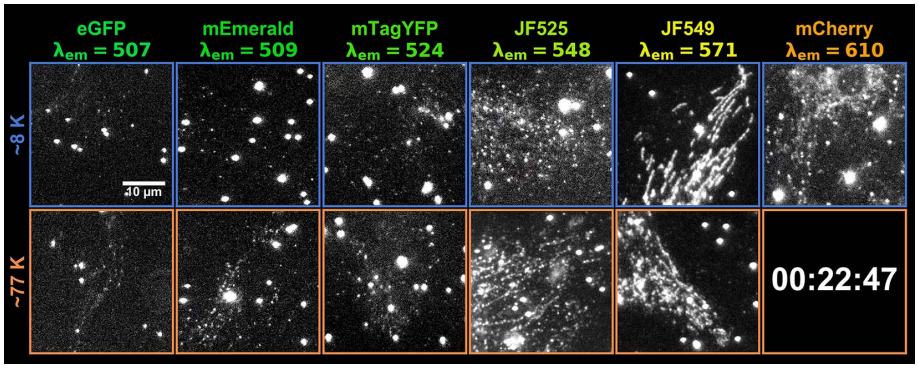
Within cells, the spatial compartmentalization of thousands of distinct proteins serves a multitude of diverse biochemical needs. Correlative super-resolution (SR) fluorescence and electron microscopy (EM) can elucidate protein spatial relationships to global ultrastructure, but has suffered from tradeoffs of structure preservation, fluorescence retention, resolution, and field of view. We developed a platform for three-dimensional cryogenic SR and focused ion beam–milled block-face EM across entire vitreously frozen cells. The approach preserves ultrastructure while enabling independent SR and EM workflow optimization. We discovered unexpected protein-ultrastructure relationships in mammalian cells including intranuclear vesicles containing endoplasmic reticulum–associated proteins, web-like adhesions between cultured neurons, and chromatin domains subclassified on the basis of transcriptional activity. Our findings illustrate the value of a comprehensive multimodal view of ultrastructural variability across whole cells.

Electron microscopy (EM) has revealed an intricate world inside eukaryotic cells (*1*), spatially organized at all length scales from nanometer-sized molecular assemblies to cell-spanning structures such as actin stress fibers and microtubules. However, even within different regions of the cell, there are notable differences in the structure of individual components, such as nuclear chromatin organization (*2*) or the morphology of the endoplasmic reticulum (ER), which is highly convoluted and compact in the perinuclear region, yet sparsely reticulated in lamellipodia (*1*). Thus, a comprehensive picture of cellular organization requires nanometer-level three-dimensional (3D) imaging of whole cells.

Although cryogenic EM (cryo-EM) tomography offers subnanometer 3D resolution (*3*), it is limited to sparse deposits of extracted macromolecules, cellular sections of submicrometer thickness (*4*–*7*), or thin lamella sculpted with cryo–focused ion beam (FIB) milling (*8*, *9*). In contrast, serial FIB ablation and imaging of the exposed face of resin-embedded specimens by scanning electron microscopy (FIB-SEM) routinely achieves 8-nm isotropic 3D sampling (*10*–*12*), a degree of precision not possible with traditional 3D EM by diamond knife serial array (*13*, *14*) or block-face sectioning (*15*). However, EM produces grayscale images in which the unambiguous identification and 3D segmentation of many subcellular structures can be challenging, and where the distributions of specific proteins can rarely be identified.

In response, correlative light and electron microscopy (CLEM) techniques have been developed that combine the global contrast and high resolution of EM with the molecular specificity of fluorescence microscopy (*16*, *17*). With the advent of super-resolution (SR) microscopy (*18*), such techniques now offer a closer match in resolution between the two modalities (table S1 and text S1), allowing specific molecular components to be visualized at nanoscale resolution in the context of the crowded intracellular environment. However, SR/EM correlation often involves trade-offs in sample preparation among retention of fluorescent labels, sufficiently dense heavy metal staining for high-contrast EM, and faithful preservation of ultrastructure, particularly when chemical fixation is used (*19*–*22*).

Here, we describe a pipeline (fig. S1) for correlative cryo-SR/FIB-SEM imaging of whole cells designed to address these issues. Specifically, cryogenic, as opposed to room-temperature, SR performed after high-pressure freezing (HPF) allowed us to use a standard EM sample preparation protocol without compromise. We correlated cryogenic 3D structured illumination microscopy (SIM) and single-molecule localization microscopy (SMLM) SR image volumes, revealing protein-specific contrast with 3D FIB-SEM image volumes containing global contrast of subcellular ultrastructure. The SR



Movie 1. Raw single-molecule frames over time since initial illumination, illustrating dark-state conversion efficiency and background as functions of temperature and emission wavelength. High pressure–frozen U2OS cells expressing fluorescent protein or dye-labeled TOMM20 to mark the outer mitochondrial membrane are shown at 10 different intervals over 3.5 hours of illumination. Bright continuous emitters are fluorescent-bead fiducial markers. As seen, all six emitters exhibit better single-molecule contrast at ~8 K than at 77 K, yielding more accurate single-molecule localization (Fig. 1, A to C, and fig. S9).

[1]Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA. [2]Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [3]Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA. [4]Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA 02115, USA. [5]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA. [6]Neuroimaging Laboratory, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [7]Bioimage Analysis Core, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [8]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA. [9]Department of Physics, University of California, Berkeley, CA 94720, USA. [10]Howard Hughes Medical Institute, Berkeley, CA 94720, USA. [11]Helen Wills Neuroscience Institute, Berkeley, CA 94720, USA. [12]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.
*These authors contributed equally to this work. †Present address: Alnylam Pharmaceuticals, 300 Third Street, Cambridge, MA 02142, USA. ‡Present address: Electron Microscopy Resource Center, The Rockefeller University, New York, NY 10065, USA. §Present address: Division of Gene Regulation, Netherlands Cancer Institute, Amsterdam 1066 CX, Netherlands.
#Corresponding author. Email: hessh@janelia.hhmi.org (H.F.H.); betzige@janelia.hhmi.org (E.B.)

modality highlights features not readily apparent from the EM data alone, such as exceptionally long or convoluted endosomes, and permits unique classification of vesicles of similar morphology such as lysosomes, peroxisomes, and mitochondrial-derived vesicles. Cell-wide 3D correlation also reveals unexpected localization patterns of proteins, including intranuclear vesicles positive for an ER marker, intricate web-like structures of adhesion proteins at cell-cell junctions, and heterogeneity in euchromatin or heterochromatin recruitment of transcriptionally associated histone H3.3 and heterochromatin protein 1α (HP1α) in the nuclei of neural progenitor cells as they transition into differentiated neurons. More generally, whole-cell cryo-SR/FIB-SEM can reveal compartmentalized proteins within known subcellular components, aid in the discovery of new subcellular components, and classify unknown EM morphologies and their roles in cell biology.

Cryogenic SR below 10 K: Motivations and photophysical characterization

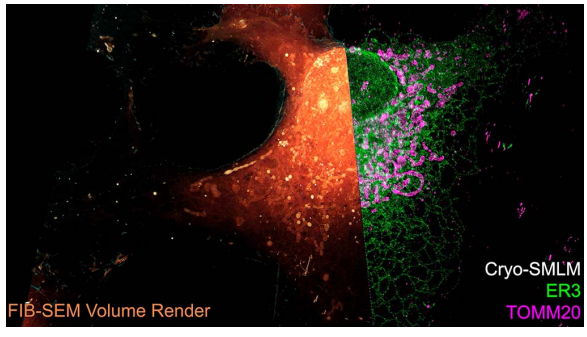
To avoid artifacts associated with chemical fixation (fig. S2), our pipeline begins with cryo-fixation via HPF (*23*, *24*) of whole cells cultured on sapphire disks 3 mm in diameter and 50 μm thick (text S2). Unlike plunge-freeze methods, HPF reliably freezes specimens up to 200 μm thick (*21*, *23*, *25*, *26*) in their entirety within vitreous ice in milliseconds, providing an exact snapshot of subcellular ultrastructure (fig. S3 and movie S1). Each sapphire disk provides an optically flat and transparent back surface for aberration-free SR imaging, along with the high thermal conductivity needed to minimize specimen heating and potential ice recrystallization under the intense (~kW/cm^2), long-lasting illumination used during SMLM. Frozen specimens are inspected, cleaned (movie S2), and loaded onto a solid copper sample holder (fig. S4) in a covered, liquid nitrogen (LN_2)–cooled preparation chamber (fig. S5) before transfer through a load lock to an evacuated optical cryostat modified for SR imaging (fig. S6).

Cryo-SR increases fluorophore photostability (*27*). This allowed us to achieve the high photon counts required for precise single-molecule localization, despite the modest numerical aperture (NA 0.85) we were compelled to use in order to image through the cryostat window, vacuum, and sapphire substrate (fig. S1 and text S3). This, along with a high reactivation efficiency under 405-nm illumination (*28*, *29*), allowed us to acquire multicolor SIM/SMLM images of the same cells without substantial photobleaching. In turn, this enabled SIM/SMLM correlation in three or more colors (movie S3) and allowed us to quickly image and assess many cells across the substrate by 3D SIM. We could then concentrate on the best candidates for much slower, higher-resolution imaging by 3D SMLM.

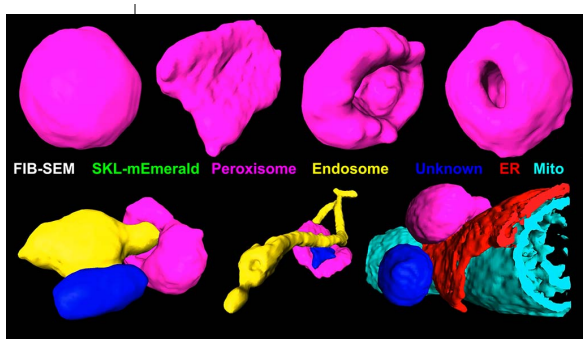
Most cryo-SR systems to date operate with LN_2 cooling near 77 K (*7*, *27*, *29*–*34*). However, we opted for a liquid helium (LHe)–cooled microscope, which allowed us to explore photophysics at any temperature down to 8 K (text S4). In particular, we exploited a sharp increase in the lifetime of a dark state D_1 for many fluorescent molecules with decreasing temperature (fig. S7) that allowed them to be shelved efficiently for long periods. Such shelving has important implications for SMLM because it dictates the dynamic contrast ratio (DCR), defined by the time a given molecule is OFF and shelved in the dark state normalized to the time it is ON and cycling between singlet states S_0 and S_1 (fig. S7) to emit light. Molecules with high DCR can be expressed at higher density, creating SMLM images of higher fidelity and resolution, with less chance of spontaneous overlap of the diffraction spots from multiple molecules that would otherwise hinder precise localization.

We measured (Fig. 1A) the DCR of six different fluorophores at both 8 K and 77 K from the ON/OFF blinking behavior of isolated single molecules (fig. S8 and text S4a). In addition, we compared (Fig. 1B) their static contrast ratios (SCR, defined by the ratio of their signal in the ON state to their local background; fig. S9), which must also be high for precise localization, during SMLM imaging of densely labeled mitochondria (Movie 1, fig. S9, and text S4b). DCR and SCR tended to increase with shorter emission wavelengths, making such fluorophores better suited to high-quality SMLM imaging (Fig. 1C). SCR also often improved at lower temperature (Fig. 1B). These trends are consistent with the photophysical argument that the dark-state lifetime should increase with increasing energy from D_1 to S_0, normalized to the thermal energy. In particular, we observed substantial gains in the SCR and DCR of JF525 (*35*) when operating with LHe, which, in conjunction with mEmerald, enabled high-quality two-color SMLM of densely labeled structures. However, if only cryo-SIM and/or single-color cryo-SMLM is needed, or if further study uncovers fluorophores spectrally distinct from mEmerald that work just as well at 77 K, then operation with LN_2 may prove sufficient.

To compare the relative merits of these labels for cryo-SMLM, we imaged two U2OS cells, targeting the ER membrane with mEmerald (green) and the mitochondrial outer membrane with Halo-JF525 (magenta) (*36*) or vice versa (Fig. 1D). Although both labels produced high-density, high-precision SMLM images of both targets, the Halo-JF525 images exhibited numerous bright puncta in both cases (Fig. 1D and fig. S10B). Although these may result from aggregation of Halo-tagged proteins, the presence of similar puncta in cryo-SMLM images of the ER obtained via SNAP (*37*) or CLIP (*38*) tag targeting of JF525 (fig. S10, C and D) suggest that they arise from a subset of extremely long-lived JF525 molecules that undergo numerous switching cycles. Indeed, the long persistence of JF525 and other labels at



Movie 2. Correlative cryogenic 3D super-resolution and block-face electron microscopy of whole vitreously frozen cells. Two COS-7 cells expressing markers for the endoplasmic reticulum (mEmerald-ER3, green) and mitochondria (Halo/JF525-TOMM20, magenta) are shown in relation to orthoslice (grayscale) or volume-rendered (cyan, plasma membrane; orange, intracellular volume) FIB-SEM data. An ER3-positive intranuclear vesicle and several cytosolic TOMM20-positive vesicles identified by correlation are also highlighted (Fig. 3 and figs. S15 and S16).



Movie 3. Structural diversity of peroxisomes and their interorganelle contacts. Peroxisomes from a HeLa cell expressing mEmerald-SKL are shown. Part 1: Orthoslices of the FIB-SEM and cryo-SMLM data followed by segmentations of SKL-labeled peroxisomes. Part 2: The same but with segmentations of other organelles in contact with SKL-labeled peroxisomes (Fig. 4).

8 K necessitated a fresh, data-driven approach (fig. S11 and text S5b) to the problem of correctly assigning and integrating the multiple photon bursts from each molecule. Even so, in all cases mEmerald yielded images probably more reflective of the true molecular distribution.

Cell-wide 3D correlation of cryo-SR with FIB-SEM

A key advantage of our pipeline is that inserting the cryo-SR step between cryofixation and freeze substitution/staining for FIB-SEM allowed us to decouple the sample preparation protocols for the two imaging modalities. This avoids the trade-offs of fluorescence retention, dense heavy metal staining, and ultrastructure preservation of resin embedding–based protocols (*39*–*42*). Furthermore, it allowed us to rapidly (~20 min) survey hundreds of cells across the sapphire disk (fig. S12A), so as to identify those of promising morphology and expression levels. We could then inspect these further at higher resolution by 3D cryo-SIM (5 min per cell per color) and select the very best ones of these for the ~1 to 2 days per cell per color required for 3D cryo-SMLM and ~10 to 15 days per cell needed for EM sample preparation and FIB-SEM imaging.

Thus, after cryo-SR imaging, we removed the frozen, disk-mounted specimens from the cryostat (fig. S12B) and processed them (text S6)

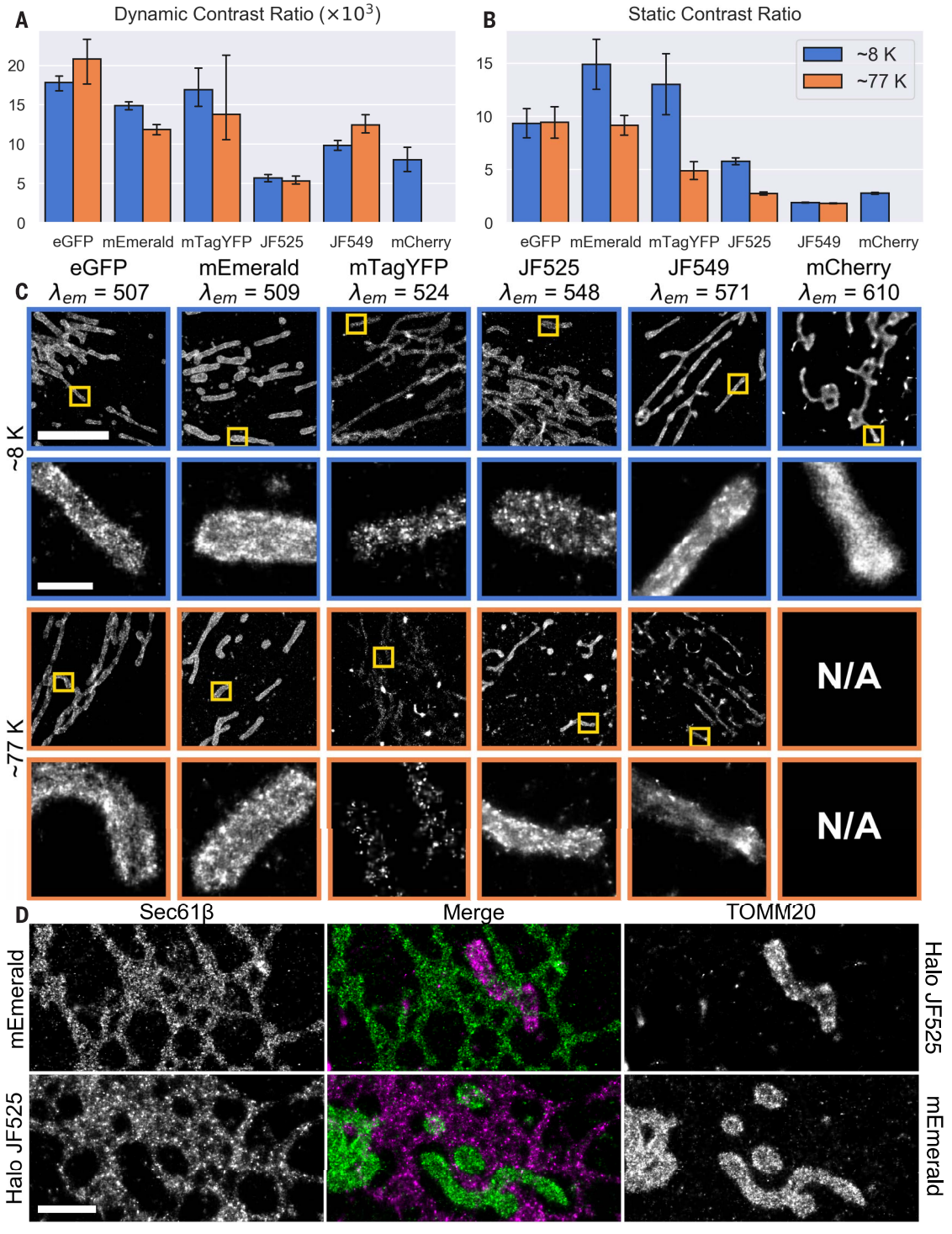


Fig. 1. Cryogenic photophysical characterization of fluorophores for single-molecule localization microscopy (SMLM). (**A** and **B**) Dynamic (A) and static (B) contrast ratios for six different fluorophores at ~8 K (blue) and ~77 K (orange) ordered by increasing emission wavelength. (**C**) Corresponding images of mitochondrial outer membrane protein TOMM20; rows 2 and 4 (scale bar, 0.5 μm) are enlargements of boxed areas in rows 1 and 3 (scale bar, 5 μm). (**D**) Comparison of mEmerald and JF525 in cryo-SMLM imaging. Top row: U2OS cell transiently expressing ER membrane marker mEmerald-Sec61β and mitochondrial membrane marker Halo-TOMM20 conjugated to JF525. Bottom row: U2OS cell transiently expressing mEmerald-TOMM20 and Halo/JF525-Sec61β. Scale bar, 1 μm.

via freeze substitution (*19*–*21*), heavy metal staining, and embedding in Eponate 12 resin (fig. S12C). After coarse trimming of the resin block and removal of the sapphire disk, we re-embedded the remaining tab in Durcupan, imaged it with x-rays (fig. S12D), and correlated this to the original disk-wide view (fig. S12E) to identify the region (typically 100 μm × 100 μm × 10 μm) containing the cells of interest imaged previously with cryo-SR. Additional microtome trimming isolated this region (fig. S12, F to I), which we then imaged at 4- to 8-nm isotropic voxels in 3D by FIB-SEM (*12*).

To exploit the full potential of correlative microscopy, the different imaging modalities need to be mutually registered to the level of their spatial resolution. Given the high resolution of both cryo-SMLM and FIB-SEM, and our desire to extend their correlation across whole cells in 3D, registration to this level is challenging. For example, slight magnification differences or deviations from ideal flat-field and rectilinear imaging coupled with potentially nonuniform FIB milling increase registration errors quickly with increasing field of view. Furthermore, freeze substitution and resin embedding introduce nonlinear and spatially inhomogeneous sample deformations (arrows, Fig. 2A) (*43*) between the cryo-SR and FIB-SEM imaging steps that have a substantial nonlinear component requiring deformable registration to achieve alignment to this level of accuracy.

Taking advantage of our protocol, we densely labeled ubiquitous intracellular organelles such as the ER and mitochondria that could be readily identified in both the cryo-SMLM and FIB-SEM data and used them as landmarks (e.g., Fig. 2B and fig. S13) to register the EM to the SR across the cellular volume (text S7), using the software package BigWarp (*44*). To quantify the accuracy of this correlation, we independently measured the deformation fields $\mathbf{DF}_{ER}$ and $\mathbf{DF}_{mito}$ from only ER or mitochondrial landmarks, respectively, after aligning these color channels to one another using fluorescent bead fiducial markers. Because $\mathbf{DF}_{ER}$ and $\mathbf{DF}_{mito}$ represent independent estimates of the underlying sample deformation, the correlation accuracy ε is given by $|\mathbf{DF}_{ER} - \mathbf{DF}_{mito}|/\sqrt{2}$ (text S7). Over a field of view covering the majority of two cells (pink surface, Fig. 2A), we measured a median ε of 89 nm (Fig. 2C), whereas in a small peripheral region (red box, Fig. 2A), the median ε was 27 nm. The difference in ε between these two regions of the same sample may be attributable to a higher density of landmarks within the peripheral region, tighter physical constraint on the differential motion between organelles due to the thinness of this region, or greater accuracy in landmark displacement measurement when the sample thickness becomes less than the axial localization precision. In any case, spatial maps of ε (fig. S14) give a local estimate of the length scale to which spatial relationships between features seen by the two imaging modalities can be reliably inferred.

Correlative cryo-SR/FIB-SEM reveals vesicle identities and their morphological diversity

Using our correlative pipeline, we imaged two neighboring COS-7 cells (Fig. 3 and Movie 2) transiently expressing mEmerald-ER3, a luminal ER marker, and HaloTag-TOMM20 conjugated to JF525 to label the mitochondrial outer membrane. Both the resulting volume rendering (Fig. 3A) and axial or transverse orthoslices (Fig. 3, B to M, and fig. S15) demonstrate accurate 3D correlation of the cryo-SMLM and FIB-SEM data, high labeling density, and faithful ultrastructure preservation throughout the ~5000-μm^3 cellular volume within the field of view.

The data also immediately illustrate the power of cryo-SR and FIB-SEM correlation. For example, clusters of ER3 seen by cryo-SMLM to dot ER tubules (orange arrows, Fig. 3B) might easily be dismissed as artifacts of labeling or fixation, but instead correlate (Fig. 3D) with varicosities in these tubules as seen by FIB-SEM (Fig. 3C). It is also immediately apparent by FIB-SEM that vesicles of various sizes are ubiquitous throughout the cell. However, these can come in many forms: peroxisomes, lysosomes, endosomes, or, as identified by our correlation, TOMM20-positive vesicles (red arrows, Fig. 3, H and J, and fig. S16). Given their small (~100 to 200 nm) size and proximity to mitochondria, these may represent mitochondria-derived vesicles (MDVs). MDVs are

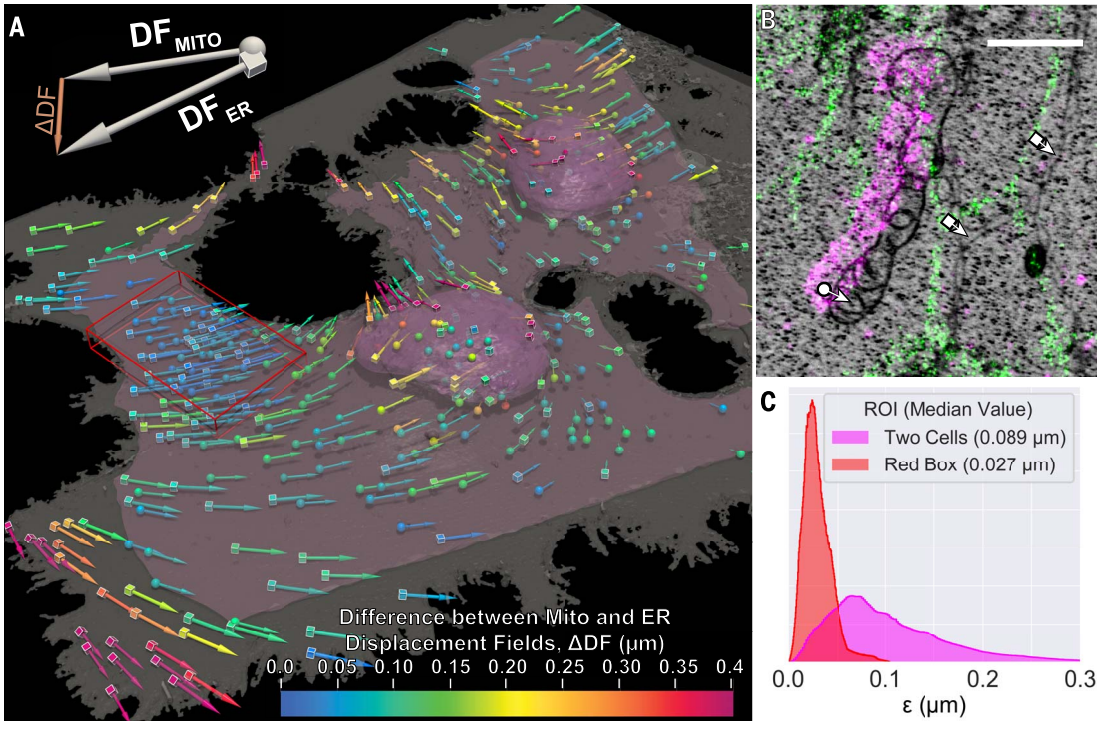


Fig. 2. High-accuracy correlation of cryo-SMLM and FIB-SEM data using organelle landmarks. (**A**) A perspective view of mitochondrial (spheres) and ER (cubes) landmarks used for registration along with the plasma (gray) and nuclear (pink) membranes as determined by FIB-SEM of two COS-7 cells. Arrows point in the direction of, and are sized according to, the magnitude of the nonlinear component of the final displacement field. Arrows are color-coded according to the magnitude of the local difference (ΔDF) between the displacement fields determined by the mitochondrial or ER landmarks separately. The pink surface indicates the boundaries of the subvolume containing sufficient landmarks of both types for quantitative comparison of their respective displacement fields. (**B**) XY orthoslice illustrating landmark selection and determination of the displacement vectors (fig. S14). Scale bar, 1 μm. (**C**) Histograms of the correlation accuracy ε (see text S7) for the subvolume bounded by the pink surface (magenta) and the 61-μm^3 subvolume defined by the red box in (A) (red), where density of both types of landmarks is higher.

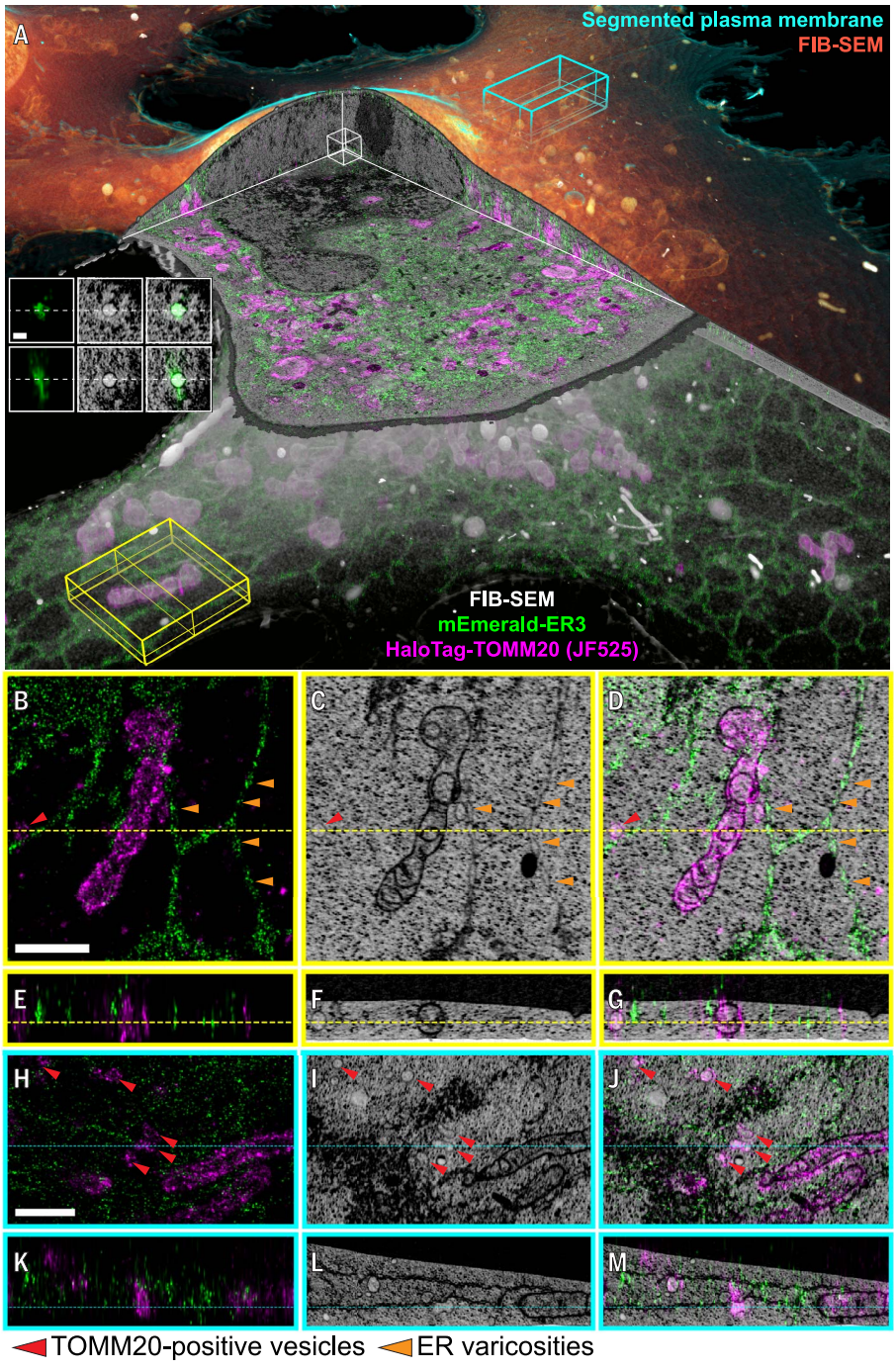


Fig. 3. Whole-cell correlative cryogenic single-molecule localization and block-face electron microscopy. (**A**) Perspective overview of a cryo-SMLM and FIB-SEM (orange and gray) dataset of a COS-7 cell transiently expressing mEmerald-ER3 (ER lumen marker, green) and Halo/JF525-TOMM20 (mitochondrial outer membrane marker, magenta) (Movie 2). Cyan, yellow, and white boxes indicate regions with orthoslices shown in (B) to (M) and inset. Inset: SMLM (left column), FIB-SEM (middle column), and correlative (right column) orthoslices in XY (top row) and XZ (bottom row) through an intranuclear ER-positive vesicle. Scale bar, 200 nm. (**B** to **G**) SMLM [(B) and (E)], FIB-SEM [(C) and (F)], and correlated overlay of orthoslices [(D) and (G)] in XY [(B) to (D)] and XZ [(E) to (G)] in a lamellipodial region. Scale bar, 1 μm. (**H** to **M**) SMLM [(H) and (K)], FIB-SEM [(I) and (L)], and correlated overlay of orthoslices [(J) and (M)] in XY [(H) to (J)] and XZ [(K) to (M)] in a thicker region with ER sheets. Scale bar, 1 μm. Red arrowheads, TOMM20-positive vesicles; orange arrowheads, varicosities in the ER.

believed to play a key role in mitochondrial quality control by sequestering unfolded or oxidized mitochondrial proteins and transporting them to lysosomes or peroxisomes for degradation (*45*). There remain, however, many questions about what proteins regulate these processes and where they are distributed.

Our data also revealed three instances of intranuclear vesicles, again ~100 to 200 nm in size, positive for the ER lumen protein ER3 (Fig. 3A, left inset orthoslices; fig. S15, correlation examples 119 and 164). In dividing somatic cells, the nuclear membrane breaks down at prometaphase and its proteins are dispersed within the ER, which remains continuous throughout mitosis (*46*). The nuclear membrane then begins to reassemble in anaphase when ER-like cisternae contact the chromatids of the daughter cells, and nuclear membrane proteins become immobilized there (*46*–*49*). Thus, one possibility is that ER lumen–positive intranuclear vesicles in interphase are the remnants of such contacts that do not completely return to the extranuclear ER after the nuclear membrane is fully reestablished. Alternatively, a small fraction of the total ER volume might be disrupted into vesicles during its rearrangement in mitosis and become similarly trapped when the nuclear membrane reforms.

Another important class of vesicles in cells are peroxisomes, which catabolize long-chain fatty acids via β-oxidation and reduce reactive oxygen species such as hydrogen peroxide (*50*). Peroxisomes can adopt a variety of sizes, shapes, and distributions depending on cell type and environment (*50*, *51*). Accurately capturing these morphologies and their spatial relationship to other organelles can be difficult with traditional chemical fixation and EM staining protocols against their enzymatic contents (*52*, *53*). Furthermore, serial section transmission EM or mechanically sectioned block-face EM (*54*) lack the axial resolution to precisely measure morphological parameters at the sub–100-nm level, whereas cryo–FIB-SEM (*55*) or cryo-EM tomography lacks the field of view to explore more than a small fraction of the total cellular volume.

We used cryo-SMLM/FIB-SEM to image and semi-automatically segment (text S8) 466 mature peroxisomes across two entire vitreously frozen HeLa cells expressing mEmerald tagged to the peroxisomal targeting sequence SKL, and Halo/JF525-TOMM20 to mark mitochondria (Fig. 4, Movie 3, and fig. S17). Independent two-channel SR/EM registration revealed a correlation accuracy (median ε) of 85 nm (fig. S14, C and D). Peroxisomes with volumes smaller than 0.01 μm^3 always assumed nearly spherical shapes, presumably to minimize their surface area under the influence of surface tension (e.g., Fig. 4, A and H). Increasingly irregular shapes such as plates (Fig. 4B), cups

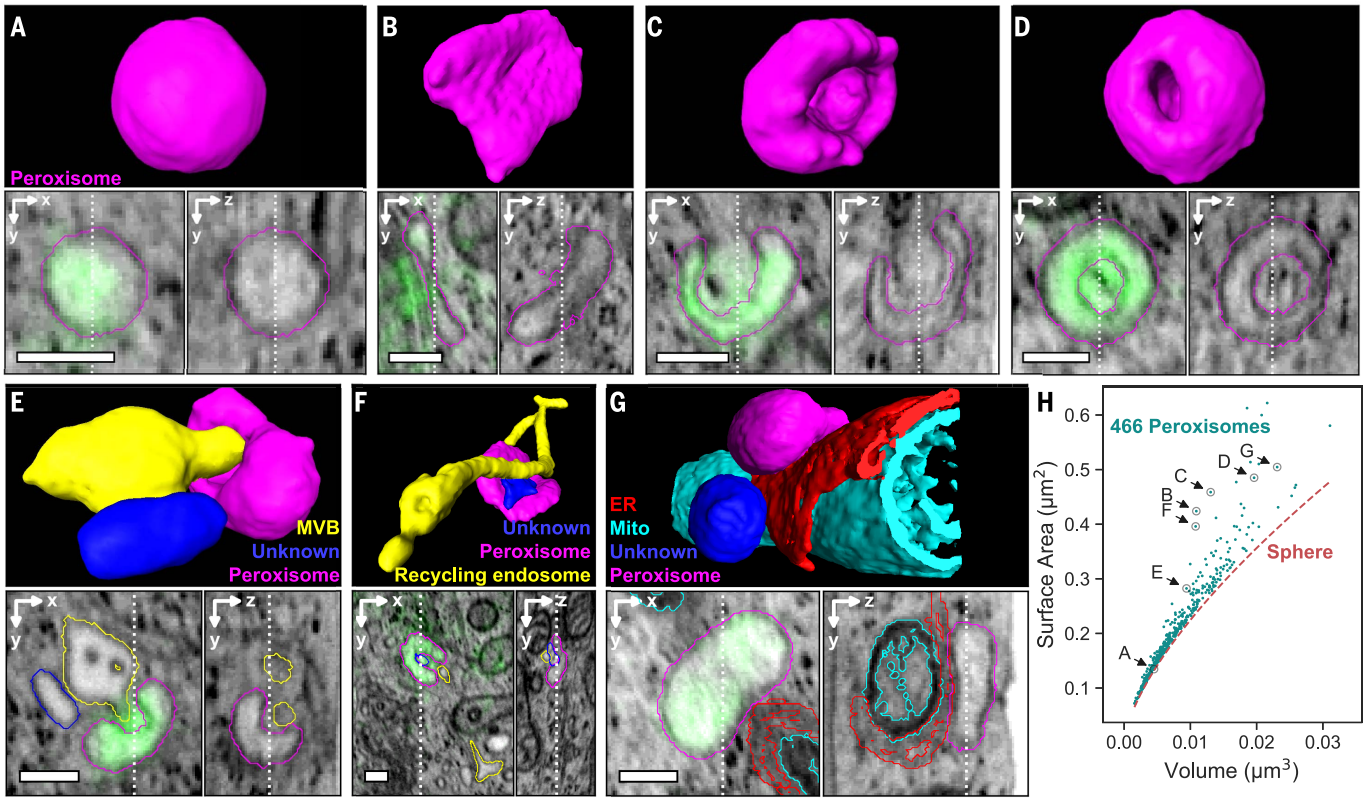
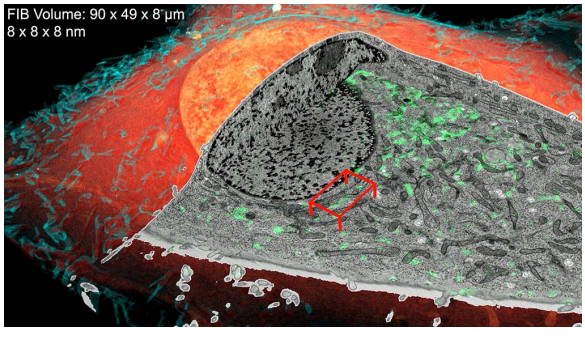


Fig. 4. Diversity of peroxisome morphologies and peroxisome-organelle interactions. (**A** to **D**) FIB-SEM segmentations (top row) of four peroxisomal targeting signal (SKL)–containing peroxisomes (magenta) and corresponding orthoslices (bottom row) with cryo-SMLM overlays of SKL (green) from two HeLa cells expressing mEmerald-SKL. (**E** to **G**) Three examples of peroxisome-organelle interactions, showing segmentations (top row) and orthoslices (bottom row) with overlaid contours of matching colors. All scale bars, 200 nm. (**H**) Surface-to-volume relationship for 466 peroxisomes (fig. S17), with the specific peroxisomes in (A) to (G) indicated, showing the increasing deviation from spherical shape with increasing volume. See also Movie 3.

(Fig. 4C), or hollow spheres (Fig. 4D) formed with increasing volume (fig. S17, lower rows). This may serve to regulate reaction kinetics rates within peroxisomes (*56*). Some irregularly shaped peroxisomes were in close proximity to other organelles or as part of multi-organelle assemblies (Fig. 4, E to G), consistent with 3D observations in live cells (*57*). These assemblies may facilitate the transfer of cargo between organelles responsible for distinct and possibly incompatible biochemical processes (*57*), such as the sequential breakdown of fatty acids between peroxisomes and mitochondria (*58*).

Finally, we explored the endolysosomal pathway, the compartments of which are notoriously sensitive to artifacts of fixation or protein overexpression (*59*–*61*). We used correlative cryo-SIM/FIB-SEM to image transferrin (Tfn)–containing endolysosomal compartments in a SUM-159 cell previously incubated for 30 min in media containing Alexa Fluor 647–conjugated Tfn (Fig. 5A and beginning of Movie 4). The density of labeled compartments was low enough to assign each discrete SIM feature (Fig. 5A, inset) to a single structure as seen by FIB-SEM, and then render each such compartment with 8-nm isotropic voxels. Despite its much lower resolution, SIM was essential to identify which compartments in the FIB-SEM data represented endolysosomes and to spot the many such structures of extremely convoluted morphology in the crowded intracellular environment that were not readily apparent by FIB-SEM alone. These included elongated tubules (Fig. 5, B to E, magenta) that likely represent recycling endosomes, highly corrugated endosomes (Fig. 5, B and D, right), and early endosomes with protruding tubules of 50 nm width possibly associated with retromers (*62*) of sub–50 nm width. Given that cryo-SIM is much faster than cryo-SMLM and can use a wide variety of spectrally distinct labels, it can be a broadly useful tool in its own right to guide the 3D segmentation of dense FIB-SEM data and ensure the correct identification of specific subcellular features.



Movie 4. Cryo-SIM/FIB-SEM reveals the morphological heterogeneity of the endolysosomal system. A correlative dataset of a SUM-159 cell after endosomal uptake of Alexa Fluor–conjugated transferrin is shown. Part 1: 3D cryo-SIM data (green), correlative orthoslices, and correlative volume render (cyan, plasma membrane; orange, cellular interior). Part 2: ~13-μm^3 subvolume showing segmentations of all transferrin-containing compartments. Part 3: Same, but for a different ~20-μm^3 subvolume (Fig. 5).

Molecular underpinnings of ultrastructural specialization in neuronal cell-to-cell adhesions

Cell-to-cell adhesions mediate cell migration, nucleate cell polarity, and spur communication between individual cells

in multicellular organisms (*63*, *64*). Although the molecular context and ultrastructure of cellular adhesions to rigid artificial substrates are well characterized (*65*, *66*) those between cells in complex 3D environments are not. Neuronal adhesions are crucial for brain development, playing an integral role in sorting neurons according to their maturation status (*67*, *68*), forming the laminar structure of the brain (*69*, *70*), and ultimately promoting the complex neuronal interactions that drive circuit morphogenesis (*71*). However, they have been difficult to study because they are disrupted by chemical fixation (*19*–*21*) and because 3D geometries of neuronal contacts require isotropic 3D-EM and high-resolution LM.

We used cryo-SIM to visualize transiently expressed junctional adhesion molecule (JAM)-C (*67*, *69*, *72*), a tight-junction component, fused to JF549i-conjugated SNAP (*35*, *73*), and 2x-mVenus-drebrin, a cytoplasmic actin-microtubule crosslinker protein (*74*), in cryofixed mouse cerebellar granule neurons (CGNs) (Fig. 6A and Movie 5). The JAM-C–defined adhesion between two labeled somas was not uniform at their shared membrane contact zone (Fig. 6B) but formed a web-like structure, with drebrin preferentially associated with the edges of the JAM-C regions.

To determine whether these protein distributions correlated with membrane ultrastructure at the contact zone, we imaged the same cells by FIB-SEM (Fig. 6C). The density of heavy metal staining at the plasma membrane was also nonuniform (Fig. 6D), with the densest staining correlating perfectly with JAM-C (compare Fig. 6, B, D, and G). Moreover, the densely stained plasma membrane was less curved than the electron-lucent plasma membrane. To quantify this, we segmented the plasma membrane within the contact zone into regions of high and low electron density (Fig. 6, F and I) and then calculated the curvedness (text S9) in each (Fig. 6, H and I). The low-density plasma membrane had a curvedness of 11.3 μm^{-1}, whereas the high-density, JAM-C–rich plasma membrane had a curvedness of 5.0 μm^{-1}.

Although the smooth nature of the adhesion as defined by JAM-C is expected because of the mechanical tension induced by the juxtacrine interaction (*75*–*77*), the fact that the adhesion does not comprise the whole contact area between these two cells is unexpected. Furthermore, the enrichment of drebrin in the regions adjacent to JAM-C contrasts with the laminar stacking of adhesion-associated cytoskeletal adaptor proteins found in focal or cadherin-based adhesions on glass (*65*, *66*).

Chromatin domains and their reorganization during neuronal differentiation

In addition to adhesion, CGNs provide an excellent model system to study the cell biological underpinnings of neural development, owing to their strongly stereotyped developmental programs as they differentiate from cerebellar granule neuron progenitors (GNPs) (*78*). Intrinsic to this process is the 3D structural reorganization of their nuclear chromatin domains (*79*, *80*). To explore this in detail, we first used 3D live-cell lattice light sheet microscopy (LLSM) (*81*). Flow-sorted GNPs expressing the EGFP-Atoh1 marker of the GNP state (*82*, *83*) possessed nuclei that were significantly larger than those of terminally differentiated CGNs (Fig. 7, A and B, and text S10). Moreover, longitudinal LLSM live imaging revealed that GNPs rapidly condense their nuclei to the size of CGNs while Atoh1-EGFP expression fades (Fig. 7, A and B, and movie S4).

To uncover the intricate 3D transformations in nuclear architecture that accompany nuclear condensation during GNP differentiation, we then applied cryo-SIM to image a

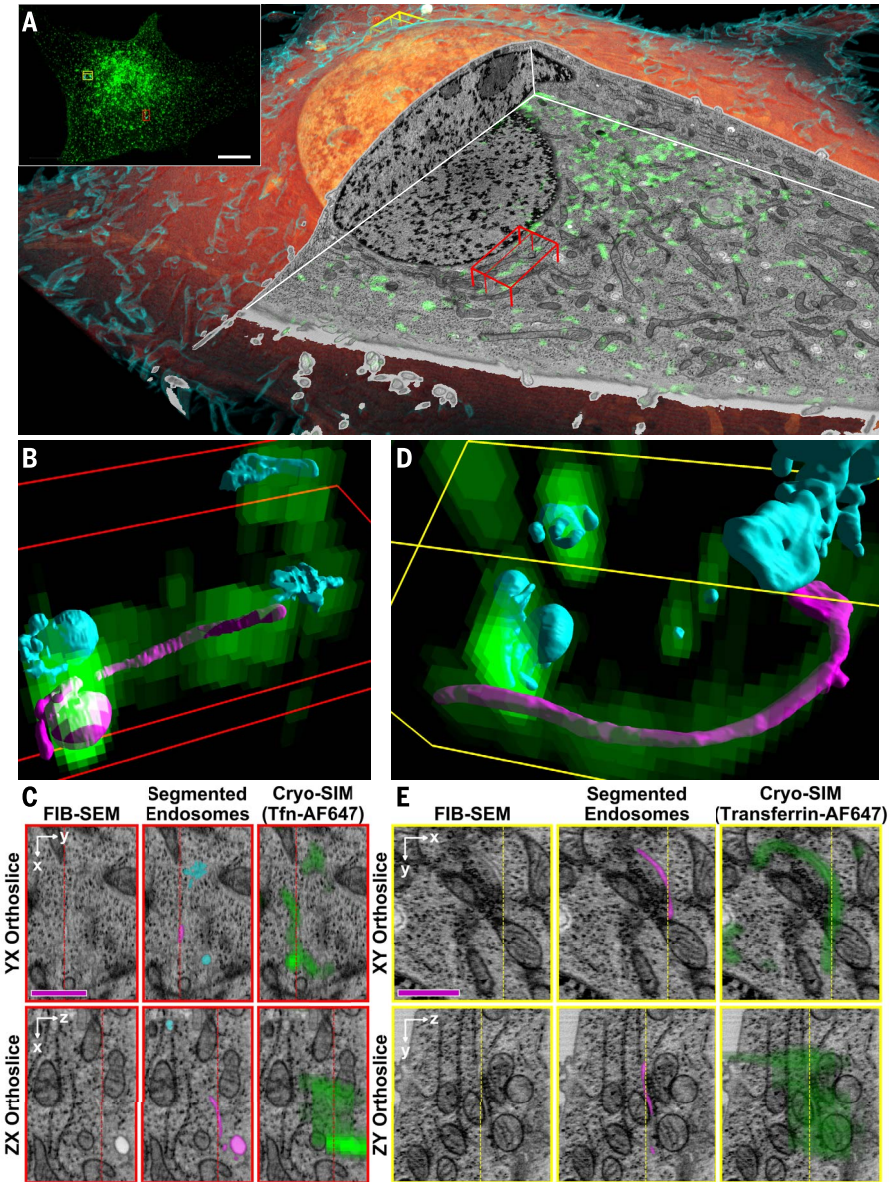


Fig. 5. Cryo-SIM/FIB-SEM accurately identifies endosomal compartments and reveals their diverse morphologies at the nanoscale. (**A**) Volume-rendered FIB-SEM overview (interior, orange; plasma membrane, cyan) of a SUM159 cell, with cutaway correlative cryo-SIM showing endolysosomal compartments containing AF647-conjugated transferrin (green). Inset: Cryo-SIM MIP of AF647-transferrin across entire field of view. (**B**) Segmented Tfn-AF647–containing compartments (colored surfaces) with superimposed 3D cryo-SIM data (green voxels) in the 13-μm^3 subvolume denoted by the red box in (A). (**C**) XY (top) and XZ (bottom) orthoslices of the same region in (B) showing FIB-SEM (left) overlaid with segmentations of transferrin-labeled compartments (middle) and cryo-SIM of Tfn-AF647 (right). (**D** and **E**) Same as (B) and (C) for the 19.5-μm^3 subvolume denoted by the yellow box in (A). Scale bars, 10 μm [(A), inset], 1 μm [(C) and (E)].

cohort of seven GNPs and nine CGNs. These collectively contained >2000 μm^3 of two nuclear domain reference proteins: mEmerald-tagged heterochromatin protein 1α (HP1α), a prototypical heterochromatin marker (*84*), and JF525-conjugated SNAP-histone 3.3 (H3.3), a replacement histone subunit that is loaded on transcriptionally active nucleosomes (*85*) (Fig. 7, C and G, top, and fig. S18). We followed this with FIB-SEM imaging and segmentation of the resulting data (*86*) according to the classic EM definitions of compacted heterochromatin, open euchromatin, and nucleoli (*2*) (Fig. 7, C and G, bottom). GNP and CGN nuclei possessed similar total nuclear volumes of compacted heterochromatin (GNP = 47 ± 2 μm^3, CGN = 45 ± 3 μm^3); however, GNP nuclei had a significantly higher total nuclear volume of euchromatin (GNP = 84 ± 8 μm^3, CGN = 61 ± 6 μm^3) that accounted for a significant fraction of the size differential with CGNs (Fig. 7K).

Registering the cryo-SIM data onto the FIB-SEM results (Fig. 7, D to F and H to J, Movie 6, figs. S19 to S21, movie S5, and text S11) allowed us to subclassify these classical EM chromatin domains according to their correlation to HP1α or H3.3 (figs. S22 and S23 and text S12). Such correlation revealed variations in these chromatin domains linked to neuronal differentiation that cannot be discerned by ultrastructure alone, including classical compacted heterochromatin domains with alternating layers HP1α and H3.3 (Fig. 7J). Indeed, although FIB-SEM showed little difference in the absolute volume of compacted heterochromatin before and after differentiation, correlation with cryo-SIM revealed that CGN nuclei had ~50% more normalized nuclear volume of HP1α-loaded heterochromatin than did GNPs (GNP = 8 ± 1%, CGN = 12 ± 2%; Fig. 7L). Moreover, measurements of surface area versus volume showed that HP1α-loaded heterochromatin became substantially more compact during nuclear condensation (fig. S23A).

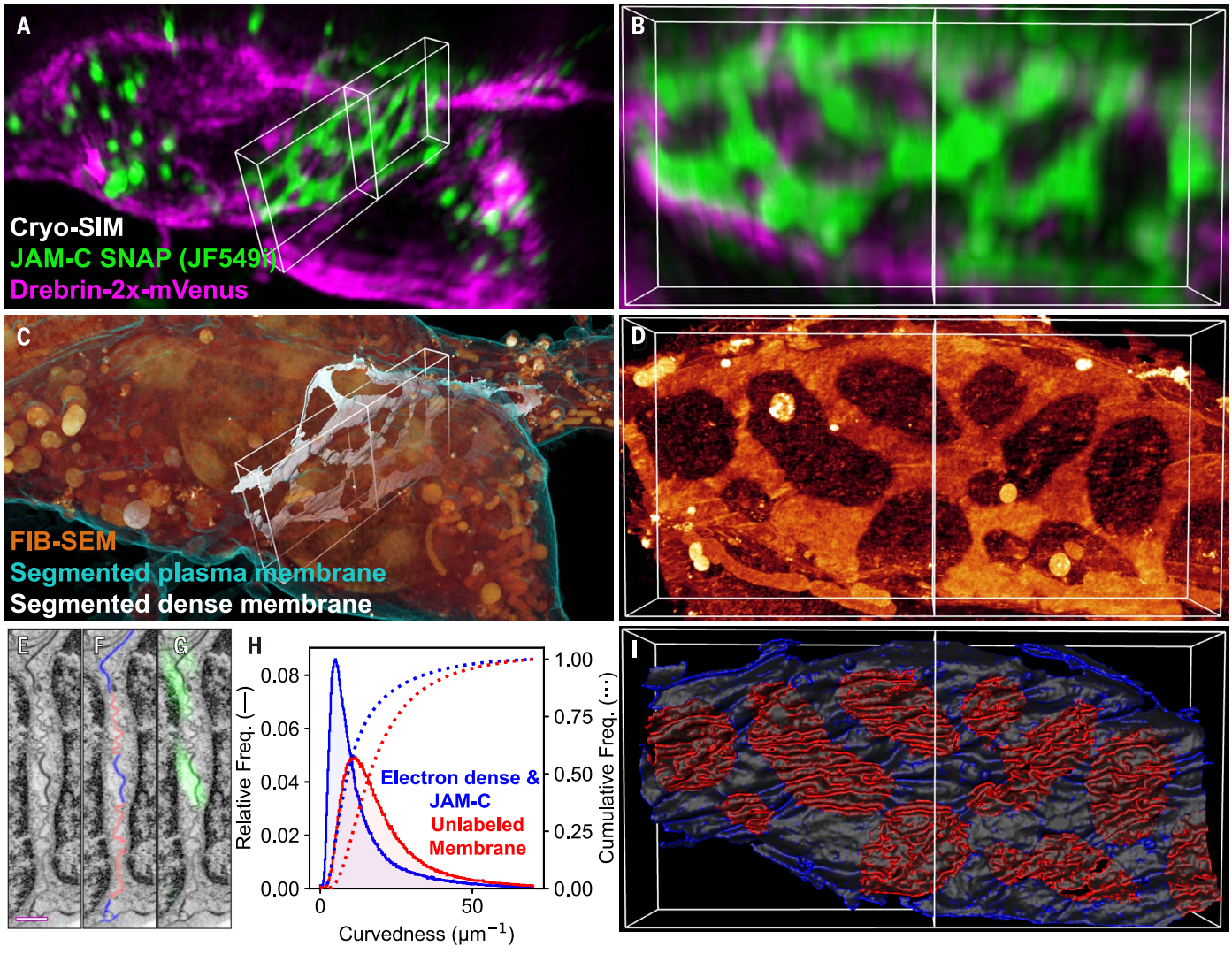


Fig. 6. Membrane proteins correlate to membrane ultrastructure at cell-cell adhesions. (**A**) Cryo-SIM volume of cultured mouse cerebellar granule neurons transiently expressing JF549i/SNAP-JAM-C (green) and 2x-mVenus-drebrin (magenta). (**B**) MIP through a slab ~3 μm thick [white box in (A)] centered on the contact zone between two cell bodies. (**C**) FIB-SEM volume of the same region in (A), with plasma membrane (cyan), intracellular content (orange), and segmented electron-dense regions of the contact zone (white). (**D**) FIB-SEM MIP through the same region in (B), after masking the nuclei. (**E**) Single FIB-SEM slice through the contact zone at the central vertical line in (D). (**F**) Same as (E), with more (blue) and less (red) electron-dense membranes traced. (**G**) Same as (E) overlaid with the JAM-C signal. Scale bar, 500 nm. (**H**) Histograms of the curvedness (text S9) for the membrane regions with high (blue) and low (red) electron density. (**I**) Partial segmentation of the cells' membranes in the contact zone, color-coded according to curvedness, with brighter colors indicating larger values. Note the high correlation among JAM-C (B), electron density (D), and membrane curvedness (I) (Movie 5). White box dimensions: 9.5 μm × 4.7 μm × 1.1 μm.

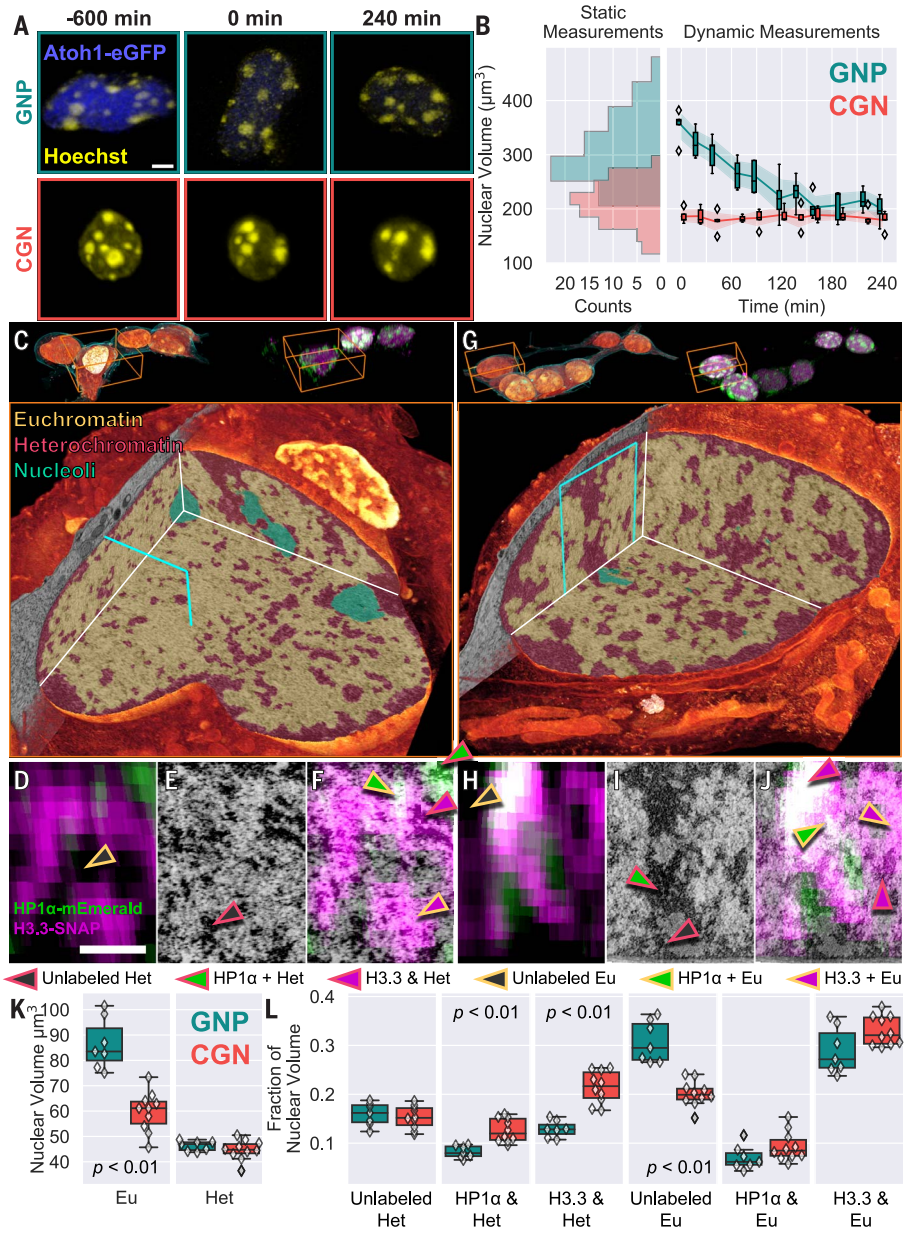


Fig. 7. Cryo-SIM/FIB-SEM reveals nuclear rearrangements associated with cerebellar granule neuron progenitor (GNP) differentiation. (**A**) Live-cell lattice light sheet time-lapse images showing an EGFP-Atoh1–positive GNP (top row) condensing its nuclear size to that of a CGN while the size of an EGFP-Atoh1–negative CGN nucleus (bottom row) remains constant. Scale bar, 3 μm. (**B**) Quantification of GNP and CGN nuclear volume for static imaging (histograms at left, 85 CGNs and 71 GNPs) and time-lapse imaging (box plots at right, 5 GNPs and 5 CGNs), showing that GNPs are on average 40% larger than CGNs and condense their nuclei to the size of CGNs in approximately 2 hours. (**C**) Top: FIB-SEM (left) and SIM (right) volume renderings of a group of GNPs. Bottom: One such GNP nucleus (orange boxes at top), with cutaway, showing color-coded chromatin territories (heterochromatin, euchromatin, or nucleoli) as identified on the basis of the EM data alone. (**D** to **F**) HP1α (green) or H3.3 (magenta) Cryo-SIM, FIB-SEM, and correlative XZ ortho slices of the plane bordered in cyan in (C). Arrowheads denote different types of labeled chromatin domains, as indicated below. Scale bar, 1 μm. (**G** to **J**) Same as (C) to (F) but for a representative CGN nucleus (Movie 6). (**K** and **L**) Quantification of EM-segmented (K) and cryo-SIM–defined (L) chromatin domains and their correlation for seven GNP and nine CGN nuclei.

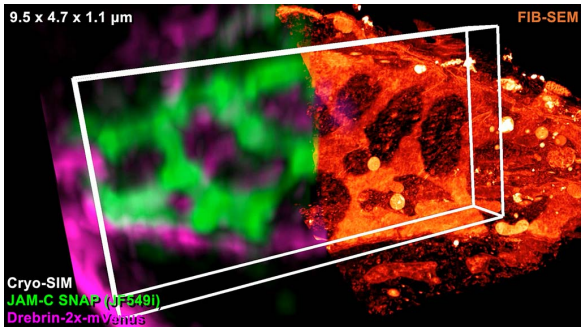
Analysis of H3.3 relationships to heterochromatin and euchromatin also revealed substantial differences between GNP and CGN nuclei. Although both GNPs and CGNs had similar amounts of H3.3-loaded euchromatin (GNP = 27 ± 4%, CGN = 32 ± 3%; Fig. 7L) indicative of transcriptionally active regions, GNPs had 50% more normalized nuclear volume of a H3.3-free form of euchromatin than did CGNs (GNP = 29 ± 4%, CGN = 20 ± 2%; Fig. 7L). Live-cell LLSM operating in the higher-resolution SIM mode revealed that these large H3.3-free voids in GNP nuclei contain mEmerald-cMAP3, a marker of H3K27me3 and H3K4me3-loaded poised chromatin (*87*), which suggests that groups of poised genes are organized in a region-specific fashion in neural progenitors (fig. S24, movie S6, and text S13).

GNP differentiation into CGNs also resulted in the unexpected accumulation of H3.3 in heterochromatin, nearly twice as abundant in CGNs as in GNPs (GNP = 13 ± 1%, CGN = 22 ± 3%; Fig. 7L). Like classical HP1α-loaded heterochromatin, H3.3-loaded heterochromatin also underwent compaction during CGN differentiation (fig. S23C). The presence of a large fraction of H3.3-loaded heterochromatin in differentiated neurons was surprising, given that H3.3-loaded heterochromatin species are abundant in pluripotent embryonic stem cells (ESCs) but have not been observed in most of their somatic cell derivatives (*88*, *89*). Furthermore, LLS-SIM revealed that H3.3-loaded heterochromatin is likely not due to H3.3 recruitment to telomeres or centromeres, as has been reported for ESCs (fig. S24B).

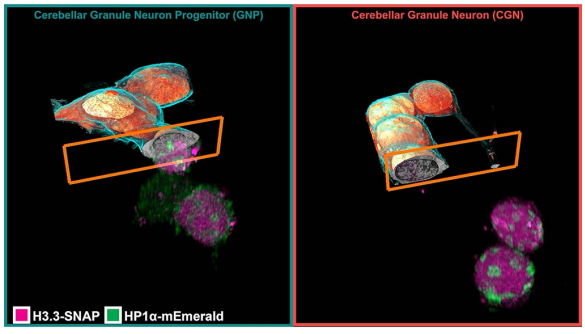
Finally, heterochromatin subdomains exhibited spatially distinct organization patterns depending on whether they were loaded with HP1α or H3.3 (fig. S22A and movie S7). Additional analysis based on the density of heavy metal staining in a correlated 4-nm FIB-SEM dataset revealed that H3.3-loaded heterochromatin was less densely packed then HP1α-loaded heterochromatin in CGN nuclei; hence, molecularly defined heterochromatin subdomains are not only spatially distinct at the level of the whole nucleus but are also morphologically distinct at the ultrastructural level (fig. S22B).

Discussion

Much of what we know about the structural and functional organization of the cell at the nanoscale comes from a synthesis of the findings of EM, biochemistry, and molecular biology. Although this synthesis has proved powerful, fusing the insights from these disparate methods necessarily involves developing models, and therefore possible biases, of how specific proteins are spatially distributed in relation to the EM ultrastructure that bear closer examination. Correlative cryo-SR/FIB-SEM enables such examination by combining two complementary datasets, often revealing unanticipated protein localization patterns or ultrastructural morphologies at variance with such models. At the same time, the approach enables the discovery of new subcategories of functionally distinct subcellular structures that appear morphologically similar when using either SR or EM alone. As such, it provides observations upon which more refined models can be developed in a



Movie 5. Correlative cryo-SIM/FIB-SEM reveals a web-like adhesion pattern between adjacent cerebellar granule neurons. Part 1: Cryo-SIM and FIB-SEM volume renderings of a field of CGNs expressing adhesion proteins JAM-C (green) and drebrin (magenta). Part 2: Correlation between electron density at the plasma membrane, JAM-C cryo-SIM signal, and plasma membrane curvature at the interface between two CGNs (Fig. 6).



Movie 6. Chromatin compaction during differentiation and identification of novel chromatin subdomains. Correlative datasets of granule neuron progenitor (GNP, left) and cerebellar granule neuron (CGN, right) are represented. Part 1: Overall correlation between the FIB-SEM (cyan, plasma membrane; orange, cellular interior) and cryo-SIM of the nuclear domain reference proteins HP1α (green) and H3.3 (magenta). Part 2: Cutaway views of EM-defined chromatin domains for a GNP nucleus (left) and a CGN nucleus (right). Part 3: Orthoslices through the CLEM volumes indicating subdomains defined by overlap between EM-defined nuclear domains and nuclear domain reference proteins. Part 4: 3D surface renderings of CLEM-defined nuclear chromatin subdomains for the GNP and CGN nuclei (Fig. 7).

way mutually consistent with the findings of SR, EM, live imaging, and biochemistry.

Of course, the value of cryo-SR/FIB-SEM to this enterprise depends on the extent to which it reveals the native ultrastructure of the cells it images, and the extent to which these cells are representative of the normal physiological state of their class. We designed our pipeline with these goals in mind. HPF immediately followed by cryo-SR imaging of cells in vitreous ice without any intervening chemical modification ensures that a faithful, unperturbed snapshot of the cell is captured, and allows SR and EM sample preparation protocols to be decoupled and independently optimized. Wide-field cryo-fluorescence imaging to rapidly survey hundreds of cells, followed by higher-resolution inspection of likely candidates by multicolor 3D cryo-SIM at a few minutes per cell, ensures that only those cells of physiological morphology, or cells in a specific desired physiological state [e.g., (*90*)], are considered for time-intensive cryo-SMLM and FIB-SEM. Lastly, freeze substitution provides excellent preservation of native ultrastructural detail while subsequent whole-cell 3D FIB-SEM gives a comprehensive picture of subcellular components across all regions of the cell, at 4- or 8-nm isotropic voxels not possible with serial-section transmission EM or mechanically sectioned serial block-face EM.

That being said, cryo-EM tomography of thin lamellae excavated from whole cells by cryo-FIB (*8*, *9*) offers molecular resolution without any risk of ultrastructural perturbation by heavy metal staining and resin embedding. Given, however, that the lamellar volume is typically only a small fraction of the entire cellular volume, many structures of interest will be missed entirely, and those that are seen may not exhibit the same morphology as in other regions of the cell. Thus, FIB-SEM and cryo-EM tomography are complementary, and developing a pipeline to do both in conjunction with cryo-SR would be a worthwhile endeavor.

Indeed, the unique ability of FIB-SEM to image whole cells and tissues at 4- to 8-nm isotropic voxels over volumes as large as 10^7 µm^3 makes it an ideal tool to map in toto the 3D ultrastructural relationships in living systems. However, to unlock its full potential, robust automated identification and segmentation of specific intracellular features of interest are required, ideally in relationship to neighboring structures with which such features might interact. This remains challenging to accomplish at scale, given the magnitude of the data involved [e.g., 100 GB in Fig. 7 and 19.5 TB in (*12*)]; the diversity, spatial density, and conformational complexity of intracellular compartments; and the monochromatic nature of the data. Cryo-SR can play an important role in the development of scalable segmentation, both in the validation of training sets for machine learning and in confirmation of the resulting segmented outputs.

We can also envision a number of possible improvements to our pipeline. First, live-cell imaging immediately prior to freezing would allow correlation of dynamics to ultrastructure (*61*), refine selection to cells of physiological behavior, and enable pharmacological, optogenetic, or other perturbations to be applied. However, the logistics for rapid and noninvasive transition from live imaging to the frozen state will require substantial technological development. Second, an extension of cryo-SR/FIB-SEM to specimens such as small gene-edited organisms or organoids that are more physiologically relevant than the isolated adherent cells with ectopically expressed markers presented here should be feasible within the 200-µm thickness limit for HPF by incorporating adaptive optics for aberration-free deep imaging. Third, the axial resolution of both cryo-SIM and cryo-SMLM could be improved by a factor of ~5 to 10 by designing a dual-window cryostat that uses opposed objectives and coherent detection, such as in I^{5}S (*91*) and iPALM (*92*). A next-generation pipeline combining these improvements could prove an even more powerful discovery platform to link 3D subcellular dynamic processes in cells, small whole organisms, and acute tissue sections to the nanoscale spatial distribution of the proteins driving these processes, all in the context of the global intracellular ultrastructure. However, even in its current form, our cryo-SR/FIB-SEM system can address a broad range of biological questions and is available to outside users wanting to do so (*93*).

Materials and methods

Preparation of vitrified samples

Specimens were cultured on 3-mm-diameter, 50-µm-thick sapphire disks (Nanjing Co-Energy Optical Crystal Co. Ltd., custom order; text S2) before cryofixation with a Wohlwend Compact 2 high-pressure freezer. See text S14 for sample-specific protocols and plasmid maps.

Cryogenic light microscopy

To optically image vitrified samples at diffraction-limited resolution and beyond, they must be maintained below 125 K to avoid devitrification (*94*) and present a clean, optically flat surface for aberration-free imaging. To achieve these ends, we built our microscope around a modified commercial liquid helium continuous-flow cryostat (Janis Research Company, ST-500; fig. S6 and text S3) and imaged cells plated on sapphire coverslips (text S2) through the opposite surface, after clearing this surface of residual ice in a custom cryo-preparation chamber (fig. S5, movie S2, and text S2). We transferred samples from cold storage to the imaging cryostat using custom tools and procedures

adapted to a commercial cryogenic vacuum transfer system (Quorum Technologies, PP3010T; fig. S6 and text S3). SIM images were processed as described (*95*); SMLM processing is described in text S15.

EM sample preparation

After optical imaging, samples were transferred back to cryo-storage before being freeze-substituted, resin-embedded, and re-embedded (text S6b and S6c). Desired regions of interest (ROIs) were identified in the plasticized specimens (fig. S12) using an XRadia 510 Versa micro X-Ray system (Carl Zeiss X-ray Microscopy Inc.) and then trimmed to expose small (~100 μm × 100 μm × 60 μm) stubs (text S6d).

FIB-SEM imaging

Standard (8 nm × 8 nm × 8 nm isotropic voxel) FIB-SEM datasets were generated using a customized Zeiss Merlin crossbeam system (*12*) and further modified as specified in text S16. The SEM image stacks were acquired at 500 kHz per pixel with an 8-nm *x-y* pixel using a 2-nA electron beam at 1.2-kV landing energy for imaging and a 15-nA gallium ion beam at 30 kV for FIB milling. Similarly, 4 nm × 4 nm × 4 nm voxel datasets were generated using a customized Zeiss GeminiSEM 500-Capella crossbeam system. The block face was imaged by a 250-pA electron beam with 0.9-kV landing energy at 200 kHz. The final image stacks were registered using a SIFT-based algorithm (*96*).

Computing resources

For most of the data analysis, except initial SMLM peak detection and fitting, we used a stand-alone Windows 10 x64 workstation with dual Xeon Gold 5122 CPUs (3.60 GHz) and 1 TB of RAM. For SMLM peak detection and fitting, we used up to 256 nodes on the Janelia cluster.

REFERENCES AND NOTES

1. D. W. Fawcett, *The Cell* (Saunders, ed. 2, 1981).
2. H. D. Ou *et al*., ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* **357**, eaag0025 (2017). doi: 10.1126/science.aag0025; pmid: 28751582
3. J. A. Briggs, Structural biology in situ—The potential of subtomogram averaging. *Curr. Opin. Struct. Biol.* **23**, 261–267 (2013). doi: 10.1016/j.sbi.2013.02.003; pmid: 23466038
4. A. Al-Amoudi *et al*., Cryo-electron microscopy of vitreous sections. *EMBO J.* **23**, 3583–3588 (2004). doi: 10.1038/sj.emboj.7600366; pmid: 15318169
5. A. Al-Amoudi, L. P. O. Norlen, J. Dubochet, Cryo-electron microscopy of vitreous sections of native biological cells and tissues. *J. Struct. Biol.* **148**, 131–135 (2004). doi: 10.1016/j.jsb.2004.03.010; pmid: 15363793
6. A. Al-Amoudi, D. Studer, J. Dubochet, Cutting artefacts and cutting process in vitreous sections for cryo-electron microscopy. *J. Struct. Biol.* **150**, 109–121 (2005). doi: 10.1016/j.jsb.2005.01.003; pmid: 15797735
7. B. Liu *et al*., Three-dimensional super-resolution protein localization correlated with vitrified cellular context. *Sci. Rep.* **5**, 13017 (2015). doi: 10.1038/srep13017; pmid: 26462878
8. M. Schaffer *et al*., Optimized cryo-focused ion beam sample preparation aimed at in situ structural studies of membrane proteins. *J. Struct. Biol.* **197**, 73–82 (2017). doi: 10.1016/j.jsb.2016.07.010; pmid: 27444390
9. S. Pfeffer, J. Mahamid, Unravelling molecular complexity in structural cell biology. *Curr. Opin. Struct. Biol.* **52**, 111–118 (2018). doi: 10.1016/j.sbi.2018.08.009; pmid: 30339965
10. G. Knott, H. Marchman, D. Wall, B. Lich, Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. *J. Neurosci.* **28**, 2959–2964 (2008). doi: 10.1523/JNEUROSCI.3189-07.2008; pmid: 18353998
11. K. Narayan, S. Subramaniam, Focused ion beams in biology. *Nat. Methods* **12**, 1021–1031 (2015). doi: 10.1038/nmeth.3623; pmid: 26513553
12. C. S. Xu *et al*., Enhanced FIB-SEM systems for large-volume 3D imaging. *eLife* **6**, e25916 (2017). doi: 10.7554/eLife.25916; pmid: 28500755
13. K. J. Hayworth, N. Kasthuri, R. Schalek, J. W. Lichtman, Automating the Collection of Ultrathin Serial Sections for Large Volume TEM Reconstructions. *Microsc. Microanal.* **12** (S02), 86–87 (2006). doi: 10.1017/S1431927606066268
14. D. D. Bock *et al*., Network anatomy and *in vivo* physiology of visual cortical neurons. *Nature* **471**, 177–182 (2011). doi: 10.1038/nature09802; pmid: 21390124
15. W. Denk, H. Horstmann, Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLOS Biol.* **2**, e329 (2004). doi: 10.1371/journal.pbio.0020329; pmid: 15514700
16. P. de Boer, J. P. Hoogenboom, B. N. G. Giepmans, Correlated light and electron microscopy: Ultrastructure lights up! *Nat. Methods* **12**, 503–513 (2015). doi: 10.1038/nmeth.3400; pmid: 26020503
17. M. Hauser *et al*., Correlative Super-Resolution Microscopy: New Dimensions and New Opportunities. *Chem. Rev.* **117**, 7428–7456 (2017). doi: 10.1021/acs.chemrev.6b00604; pmid: 28045508
18. E. Wegel *et al*., Imaging cellular structures in super-resolution with SIM, STED and Localisation Microscopy: A practical comparison. *Sci. Rep.* **6**, 27290 (2016). doi: 10.1038/srep27290; pmid: 27264341
19. D. R. Keene, K. McDonald, The ultrastructure of the connective tissue matrix of skin and cartilage after high-pressure freezing and freeze-substitution. *J. Histochem. Cytochem.* **41**, 1141–1153 (1993). doi: 10.1177/41.8.8331280; pmid: 8331280
20. K. L. McDonald, Electron microscopy and EM immunocytochemistry. *Methods Cell Biol.* **44**, 411–444 (1994). doi: 10.1038/nmeth.1855; pmid: 22290187
21. K. McDonald, M. Morphew, P. Verkade, T. Müller-Reichert, Recent advances in high-pressure freezing: Equipment and specimen loading methods. *Methods Mol. Biol.* **369**, 143–173 (2007).doi: 10.1038/nmeth.1855; pmid: 22290187
22. U. Schnell, F. Dijk, K. A. Sjollema, B. N. G. Giepmans, Immunolabeling artifacts and the need for live-cell imaging. *Nat. Methods* **9**, 152–158 (2012). doi: 10.1038/nmeth.1855; pmid: 22290187
23. D. Studer, B. M. Humbel, M. Chiquet, Electron microscopy of high pressure frozen samples: Bridging the gap between cellular ultrastructure and atomic resolution. *Histochem. Cell Biol.* **130**, 877–889 (2008). doi: 10.1007/s00418-008-0500-1; pmid: 18795316
24. D. Studer *et al*., Capture of activity-induced ultrastructural changes at synapses by high-pressure freezing of brain tissue. *Nat. Protoc.* **9**, 1480–1495 (2014). doi: 10.1038/nprot.2014.099; pmid: 24874814
25. J. Z. Kiss, T. H. Giddings Jr., L. A. Staehelin, F. D. Sack, Comparison of the ultrastructure of conventionally fixed and high pressure frozen/freeze substituted root tips of Nicotiana and Arabidopsis. *Protoplasma* **157**, 64–74 (1990). doi: 10.1007/BF01322639; pmid: 11538077
26. R. Dahl, L. A. Staehelin, High-pressure freezing for the preservation of biological structure: Theory and practice. *J. Electron Microsc. Tech.* **13**, 165–174 (1989). doi: 10.1002/jemt.1060130305; pmid: 2685196
27. W. Li, S. C. Stein, I. Gregor, J. Enderlein, Ultra-stable and versatile widefield cryo-fluorescence microscope for single-molecule localization with sub-nanometer accuracy. *Opt. Express* **23**, 3770–3783 (2015). doi: 10.1364/OE.23.003770; pmid: 25836229
28. M. A. Schwentker, thesis, Kirchhoff Institute for Physics (2007). doi: 10.11588/heidok.00007677
29. M. W. Tuijtel, A. J. Koster, S. Jakobs, F. G. A. Faas, T. H. Sharp, Correlative cryo super-resolution light and electron microscopy on mammalian cells using fluorescent proteins. *Sci. Rep.* **9**, 1369 (2019). doi: 10.1038/s41598-018-37728-8; pmid: 30718653
30. Y.-W. Chang *et al*., Correlated cryogenic photoactivated localization microscopy and cryo-electron tomography. *Nat. Methods* **11**, 737–739 (2014). doi: 10.1038/nmeth.2961; pmid: 24813625
31. R. Kaufmann *et al*., Super-resolution microscopy using standard fluorescent proteins in intact cells under cryo-conditions. *Nano Lett.* **14**, 4171–4175 (2014). doi: 10.1021/nl501870p; pmid: 24884378
32. P. D. Dahlberg *et al*., Identification of PAmKate as a Red Photoactivatable Fluorescent Protein for Cryogenic Super-Resolution Imaging. *J. Am. Chem. Soc.* **140**, 12310–12313 (2018). doi: 10.1021/jacs.8b05960; pmid: 30222332
33. F. Moser *et al*., Cryo-SOFI enabling low-dose super-resolution correlative light and electron cryo-microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4804–4809 (2019). doi: 10.1073/pnas.1810690116; pmid: 30808803
34. L. Wang *et al*., Solid immersion microscopy images cells under cryogenic conditions with 12 nm resolution. *Commun. Biol.* **2**, 74 (2019). doi: 10.1038/s42003-019-0317-6; pmid: 30820469
35. J. B. Grimm *et al*., A general method to fine-tune fluorophores for live-cell and *in vivo* imaging. *Nat. Methods* **14**, 987–994 (2017). doi: 10.1038/nmeth.4403; pmid: 28869757
36. G. V. Los *et al*., HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373–382 (2008). doi: 10.1021/cb800025k; pmid: 18533659
37. A. Keppler *et al*., A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **21**, 86–89 (2003). doi: 10.1038/nbt765; pmid: 12469133
38. A. Gautier *et al*., An engineered protein tag for multiprotein labeling in living cells. *Chem. Biol.* **15**, 128–136 (2008). doi: 10.1016/j.chembiol.2008.01.007; pmid: 18291317
39. D. Kim *et al*., Correlative stochastic optical reconstruction microscopy and electron microscopy. *PLOS ONE* **10**, e0124581 (2015). doi: 10.1371/journal.pone.0124581; pmid: 25874453
40. C. J. Peddie *et al*., Correlative super-resolution fluorescence and electron microscopy using conventional fluorescent proteins in vacuo. *J. Struct. Biol.* **199**, 120–131 (2017). doi: 10.1016/j.jsb.2017.05.013; pmid: 28576556
41. M. G. Paez-Segala *et al*., Fixation-resistant photoactivatable fluorescent proteins for CLEM. *Nat. Methods* **12**, 215–218 (2015). doi: 10.1038/nmeth.3225; pmid: 25581799
42. E. Johnson *et al*., Correlative in-resin super-resolution and electron microscopy using standard fluorescent proteins. *Sci. Rep.* **5**, 9583 (2015). doi: 10.1038/srep09583; pmid: 25823571
43. N. Matsko, M. Mueller, Epoxy resin as fixative during freeze-substitution. *J. Struct. Biol.* **152**, 92–103 (2005). doi: 10.1016/j.jsb.2005.07.005; pmid: 16214372
44. J. A. Bogovic, P. Hanslovsky, A. Wong, S. Saalfeld, in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)* (IEEE, 2016), pp. 1123–1126; http://ieeexplore.ieee.org/document/7493463/.
45. A. Sugiura, G.-L. McLelland, E. A. Fon, H. M. McBride, A new pathway for mitochondrial quality control: Mitochondrial-derived vesicles. *EMBO J.* **33**, 2142–2156 (2014). doi: 10.15252/embj.201488104; pmid: 25107473
46. B. Burke, J. Ellenberg, Remodelling the walls of the nucleus. *Nat. Rev. Mol. Cell Biol.* **3**, 487–497 (2002). doi: 10.1038/nrm860; pmid: 12094215
47. D. J. Anderson, M. W. Hetzer, Reshaping of the endoplasmic reticulum limits the rate for nuclear envelope formation. *J. Cell Biol.* **182**, 911–924 (2008). doi: 10.1083/jcb.200805140; pmid: 18779370
48. J. Ellenberg *et al*., Nuclear membrane dynamics and reassembly in living cells: Targeting of an inner nuclear membrane protein in interphase and mitosis. *J. Cell Biol.* **138**, 1193–1206 (1997). doi: 10.1083/jcb.138.6.1193; pmid: 9298976
49. L. Lu, M. S. Ladinsky, T. Kirchhausen, Cisternal organization of the endoplasmic reticulum during mitosis. *Mol. Biol. Cell* **20**, 3471–3480 (2009). doi: 10.1091/mbc.e09-04-0327; pmid: 19494040
50. J. J. Smith, J. D. Aitchison, Peroxisomes take shape. *Nat. Rev. Mol. Cell Biol.* **14**, 803–817 (2013). doi: 10.1038/nrm3700; pmid: 24263361
51. M. Grabenbauer, K. Sätzler, E. Baumgart, H. D. Fahimi, Three-dimensional ultrastructural analysis of peroxisomes in HepG2 cells. *Cell Biochem. Biophys.* **32**, 37–49 (2000). doi: 10.1385/CBB:32:1-3:37; pmid: 11330069
52. M. Veenhuis, S. E. Wendelaar Bonga, Cytochemical localization of catalase and several hydrogen peroxide-producing oxidases in the nucleoids and matrix of rat liver peroxisomes. *Histochem. J.* **11**, 561–572 (1979). doi: 10.1007/BF01012539; pmid: 511592
53. S. Angermüller, H. D. Fahimi, Selective cytochemical localization of peroxidase, cytochrome oxidase and catalase in rat liver with 3,3′-diaminobenzidine. *Histochemistry* **71**, 33–44 (1981). doi: 10.1007/BF00592568; pmid: 6262282

54. B. Titze, C. Genoud, Volume scanning electron microscopy for imaging biological ultrastructure. *Biol. Cell* **108**, 307–323 (2016). doi: 10.1111/boc.201600024; pmid: 27432264
55. N. Vidavsky *et al*., Cryo-FIB-SEM serial milling and block face imaging: Large volume structural analysis of biological tissues preserved close to their native state. *J. Struct. Biol.* **196**, 487–495 (2016). doi: 10.1016/j.jsb.2016.09.016; pmid: 27693309
56. L. Lizana, B. Bauer, O. Orwar, Controlling the rates of biochemical reactions and signaling networks by shape and volume changes. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4099–4104 (2008). doi: 10.1073/pnas.0709932105; pmid: 18337513
57. A. M. Valm *et al*., Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162–167 (2017). doi: 10.1038/nature22369; pmid: 28538724
58. B. Alberts, Ed., *Molecular Biology of the Cell* (Garland Science, ed. 4, 2002).
59. J. Klumperman, G. Raposo, The complex ultrastructure of the endolysosomal system. *Cold Spring Harb. Perspect. Biol.* **6**, a016857 (2014). doi: 10.1101/cshperspect.a016857; pmid: 24851870
60. J. R. Goldenring, Recycling endosomes. *Curr. Opin. Cell Biol.* **35**, 117–122 (2015). doi: 10.1016/j.ceb.2015.04.018; pmid: 26022676
61. J. Fermie *et al*., Single organelle dynamics linked to 3D structure by correlative live-cell imaging and 3D electron microscopy. *Traffic* **19**, 354–369 (2018). doi: 10.1111/tra.12557; pmid: 29451726
62. I. J. McGough, P. J. Cullen, Recent advances in retromer biology. *Traffic* **12**, 963–971 (2011). doi: 10.1111/j.1600-0854.2011.01201.x; pmid: 21463457
63. T. Lecuit, Adhesion remodeling underlying tissue morphogenesis. *Trends Cell Biol.* **15**, 34–42 (2005). doi: 10.1016/j.tcb.2004.11.007; pmid: 15653076
64. P. McMillen, S. A. Holley, Integration of cell-cell and cell-ECM adhesion in vertebrate morphogenesis. *Curr. Opin. Cell Biol.* **36**, 48–53 (2015). doi: 10.1016/j.ceb.2015.07.002; pmid: 26189063
65. P. Kanchanawong *et al*., Nanoscale architecture of integrin-based cell adhesions. *Nature* **468**, 580–584 (2010). doi: 10.1038/nature09621; pmid: 21107430
66. C. Bertocchi *et al*., Nanoscale architecture of cadherin-based cell adhesions. *Nat. Cell Biol.* **19**, 28–37 (2017). doi: 10.1038/ncb3456; pmid: 27992406
67. D. J. Solecki, Sticky situations: Recent advances in control of cell adhesion during neuronal migration. *Curr. Opin. Neurobiol.* **22**, 791–798 (2012). doi: 10.1016/j.conb.2012.04.010; pmid: 22560352
68. W. O. Gao, N. Heintz, M. E. Hatten, Cerebellar granule cell neurogenesis is regulated by cell-cell interactions in vitro. *Neuron* **6**, 705–715 (1991). doi: 10.1016/0896-6273(91)90168-Y; pmid: 2025426
69. J. K. Famulski *et al*., Siah regulation of Pard3A controls neuronal cell adhesion during germinal zone exit. *Science* **330**, 1834–1838 (2010). doi: 10.1126/science.1198480; pmid: 21109632
70. A. Chédotal, Should I stay or should I go? Becoming a granule cell. *Trends Neurosci.* **33**, 163–172 (2010). doi: 10.1016/j.tins.2010.01.004; pmid: 20138673
71. L. de la Torre-Ubieta, A. Bonni, Transcriptional regulation of neuronal polarity and morphogenesis in the mammalian brain. *Neuron* **72**, 22–40 (2011). doi: 10.1016/j.neuron.2011.09.018; pmid: 21982366
72. M. P. Arrate, J. M. Rodriguez, T. M. Tran, T. A. Brock, S. A. Cunningham, Cloning of human junctional adhesion molecule 3 (JAM3) and its identification as the JAM2 counter-receptor. *J. Biol. Chem.* **276**, 45826–45832 (2001). doi: 10.1074/jbc.M105972200; pmid: 11590146
73. J. B. Grimm *et al*., A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat. Methods* **12**, 244–250 (2015). doi: 10.1038/nmeth.3256; pmid: 25599551
74. N. Trivedi *et al*., Drebrin-mediated microtubule-actomyosin coupling steers cerebellar granule neuron nucleokinesis and migration pathway selection. *Nat. Commun.* **8**, 14484 (2017). doi: 10.1038/ncomms14484; pmid: 28230156
75. R. Simson *et al*., Membrane bending modulus and adhesion energy of wild-type and mutant cells of Dictyostelium lacking talin or cortexillins. *Biophys. J.* **74**, 514–522 (1998). doi: 10.1016/S0006-3495(98)77808-7; pmid: 9449351
76. K. Lam Hui, C. Wang, B. Grooman, J. Wayt, A. Upadhyaya, Membrane dynamics correlate with formation of signaling clusters during cell spreading. *Biophys. J.* **102**, 1524–1533 (2012). doi: 10.1016/j.bpj.2012.02.015; pmid: 22500752
77. D. Stabley, S. Retterer, S. Marshall, K. Salaita, Manipulating the lateral diffusion of surface-anchored EGF demonstrates that receptor clustering modulates phosphorylation levels. *Integr. Biol.* **5**, 659–668 (2013). doi: 10.1039/c3ib20239a; pmid: 23416883
78. S. G. Kuhar *et al*., Changing patterns of gene expression define four stages of cerebellar granule neuron differentiation. *Development* **117**, 97–104 (1993). pmid: 8223263
79. C. L. Frank *et al*., Regulation of chromatin accessibility and Zic binding at enhancers in the developing cerebellum. *Nat. Neurosci.* **18**, 647–656 (2015). doi: 10.1038/nn.3995; pmid: 25849986
80. X. Zhu *et al*., Role of Tet1/3 Genes and Chromatin Remodeling Genes in Cerebellar Circuit Formation. *Neuron* **89**, 100–112 (2016). doi: 10.1016/j.neuron.2015.11.030; pmid: 26711116
81. B.-C. Chen *et al*., Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998 (2014). doi: 10.1126/science.1257998; pmid: 25342811
82. N. Ben-Arie *et al*., Math1 is essential for genesis of cerebellar granule neurons. *Nature* **390**, 169–172 (1997). doi: 10.1038/36579; pmid: 9367153
83. M. F. Rose *et al*., Math1 is essential for the development of hindbrain neurons critical for perinatal breathing. *Neuron* **64**, 341–354 (2009). doi: 10.1016/j.neuron.2009.10.023; pmid: 19914183
84. L. Fanti, S. Pimpinelli, HP1: A functionally multifaceted protein. *Curr. Opin. Genet. Dev.* **18**, 169–174 (2008). doi: 10.1016/j.gde.2008.01.009; pmid: 18329871
85. S. J. Elsaesser, A. D. Goldberg, C. D. Allis, New functions for an old variant: No substitute for histone H3.3. *Curr. Opin. Genet. Dev.* **20**, 110–117 (2010). doi: 10.1016/j.gde.2010.01.003; pmid: 20153629
86. C. Sommer, C. Straehle, U. Köthe, F. A. Hamprecht, in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2011), pp. 230–233.
87. A. M.-F. Delachat *et al*., Engineered Multivalent Sensors to Detect Coexisting Histone Modifications in Living Stem Cells. *Cell Chem. Biol.* **25**, 51–56.e6 (2018). doi: 10.1016/j.chembiol.2017.10.008; pmid: 29174541
88. L. H. Wong *et al*., Histone H3.3 incorporation provides a unique and functionally essential telomeric chromatin in embryonic stem cells. *Genome Res.* **19**, 404–414 (2009). doi: 10.1101/gr.084947.108; pmid: 19196724
89. H. P. J. Voon, L. H. Wong, New players in heterochromatin silencing: Histone variant H3.3 and the ATRX/DAXX chaperone. *Nucleic Acids Res.* **44**, 1496–1501 (2016). doi: 10.1093/nar/gkw012; pmid: 26773061
90. W. Kukulski, M. Schorb, M. Kaksonen, J. A. G. Briggs, Plasma membrane reshaping during endocytosis is revealed by time-resolved electron tomography. *Cell* **150**, 508–520 (2012). doi: 10.1016/j.cell.2012.05.046; pmid: 22863005
91. L. Shao *et al*., I5S: Wide-field light microscopy with 100-nm-scale resolution in three dimensions. *Biophys. J.* **94**, 4971–4983 (2008). doi: 10.1529/biophysj.107.120352; pmid: 18326649
92. G. Shtengel *et al*., Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3125–3130 (2009). doi: 10.1073/pnas.0813131106; pmid: 19202073
93. The cryo-SR/FIB-SEM instrument, as described here, will be available at Janelia's Advanced Imaging Center: www.aicjanelia.com.
94. J. Lepault, D. Bigot, D. Studer, I. Erk, Freezing of aqueous specimens: An X-ray diffraction study. *J. Microsc.* **187**, 158–166 (1997). doi: 10.1046/j.1365-2818.1997.2170787.x
95. M. G. L. Gustafsson *et al*., Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophys. J.* **94**, 4957–4970 (2008). doi: 10.1529/biophysj.107.120345; pmid: 18326650
96. D. G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **60**, 91–110 (2004). doi: 10.1023/B:VISI.0000029664.99615.94

ACKNOWLEDGMENTS

We thank L. Lavis, W. Legant, D. Li, L. Shao, C. Ott, N. Alivodej, N. Trivedi, D. Wong, and K. Hayworth for useful discussions; the Shared Resource teams at Janelia for their skill and dedication in specimen handling and preparation; and the Janelia Experimental Technologies team for their manufacturing expertise, in particular B. Biddle and B. Bowers. We gratefully acknowledge the support of the Janelia Visitor Program. mEmerald-N1 and mEmerald-ER3 were gifts from M. Davidson (Addgene plasmid #53976; http://n2t.net/addgene:53976; RRID:Addgene_53976 and Addgene plasmid #54082; http://n2t.net/addgene:54082; RRID:Addgene_54082). **Funding:** Supported by the Howard Hughes Medical Institute (D.P.H., G.S., C.S.X., M.F., D.E.M., H.A.P., N.I., J.A.B., S.P., D.P., K.S., C.-L.C., J.L.-S., E.B., and H.F.H.); grants from Biogen (T.K., L.W., and W.P.); NIH grant R01 GM075252 and MIRA NIH award GM130386 (T.K.); and the American Lebanese Syrian Associated Charities and National Institute of Neurological Disorders and Stroke grants 1R01NS066936 and R01NS104029-02 (K.R.C., D.R.S., A.S., and D.J.S.). LLSM images were acquired at St. Jude Children's Research Hospital in the Department of Developmental Neurobiology Neuroimaging Laboratory. **Author contributions:** D.P.H., G.S., E.B., and H.F.H. supervised the project and wrote the manuscript with input from all co-authors; D.P.H. and G.S. designed and built the cryogenic optical microscope with input from E.B. and H.F.H. and performed all characterization experiments; D.P.H. and G.S. conducted all optical measurements; C.S.X. optimized and adapted the FIB-SEM instruments; C.S.X., S.P., and D.P. conducted all EM experiments; D.E.M. created the instrument control software for the cryo-microscope; J.A.B. customized BigWarp; M.F. and K.S. prepared samples for Figs. 1 to 3 and the associated supplementary material; C.-L.C. and L.W. prepared samples for Figs. 4 and 5, respectively; D.J.S. prepared all granule neuron samples and led related experiments; D.R.S., K.R.C., and D.J.S. performed LLSM and LLSM-SIM experiments; M.F., G.S., and D.P.H. froze specimens; G.S., H.A.P., and N.I. prepared all specimens for EM; D.P.H. and G.S. processed and analyzed all data except for that presented in Fig. 7 and the associated supplementary material, which was done by K.R.C., D.R.S., A.S., and D.J.S.; D.P.H. and G.S. produced all figures and movies except figs. S18 to S24, movie S5, and the last part of Movie 6, which were produced by K.R.C. and D.R.S.; W.P. assisted in aligning the data in Fig. 5; and J.L.-S. and T.K. aided the biological interpretation of the results in Figs. 4 and 5, respectively. **Competing interests:** E.B. has a financial interest in LLSM. E.B. and H.F.H. have a financial interest in SMLM. C.S.X. and H.F.H. have a pending patent application #20180218878 (assignee: Howard Hughes Medical Institute) for the enhanced FIB-SEM system used in this work. **Data and materials availability:** All data needed to evaluate our conclusions are in the main paper or the supplementary materials. The total size of the raw data used here exceeds tens of terabytes. All data used in this paper are freely available upon reasonable request to those who provide a mechanism for facile data transfer.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/eaaz5357/suppl/DC1
Supplementary Text
Figs. S1 to S49
Tables S1 and S2
Movies S1 to S7
References (*97*–*142*)

View/request a protocol for this paper from *Bio-protocol*.

18 September 2019; accepted 20 November 2019
10.1126/science.aaz5357

RESEARCH ARTICLE

MASS EXTINCTION

On impact and volcanism across the Cretaceous-Paleogene boundary

Pincelli M. Hull[1]*†, André Bornemann[2]†, Donald E. Penman[1], Michael J. Henehan[1,3], Richard D. Norris[4], Paul A. Wilson[5], Peter Blum[6], Laia Alegret[7], Sietske J. Batenburg[8], Paul R. Bown[9], Timothy J. Bralower[10], Cecile Cournede[11,12], Alexander Deutsch[13], Barbara Donner[14], Oliver Friedrich[15], Sofie Jehle[16], Hojung Kim[9], Dick Kroon[17], Peter C. Lippert[18], Dominik Loroch[13], Iris Moebius[15,19], Kazuyoshi Moriya[20], Daniel J. Peppe[21], Gregory E. Ravizza[22], Ursula Röhl[14], Jonathan D. Schueth[23], Julio Sepúlveda[24], Philip F. Sexton[25], Elizabeth C. Sibert[4,26,27], Kasia K. Śliwińska[28], Roger E. Summons[29], Ellen Thomas[1,30], Thomas Westerhold[14], Jessica H. Whiteside[5], Tatsuhiko Yamaguchi[31], James C. Zachos[32]

The cause of the end-Cretaceous mass extinction is vigorously debated, owing to the occurrence of a very large bolide impact and flood basalt volcanism near the boundary. Disentangling their relative importance is complicated by uncertainty regarding kill mechanisms and the relative timing of volcanogenic outgassing, impact, and extinction. We used carbon cycle modeling and paleotemperature records to constrain the timing of volcanogenic outgassing. We found support for major outgassing beginning and ending distinctly before the impact, with only the impact coinciding with mass extinction and biologically amplified carbon cycle change. Our models show that these extinction-related carbon cycle changes would have allowed the ocean to absorb massive amounts of carbon dioxide, thus limiting the global warming otherwise expected from postextinction volcanism.

Sixty-six million years ago, two planetary-scale disturbances occurred within less than a million years of one another. One disturbance was the collision of an asteroid of more than 10 km in diameter with the Yucatan Peninsula at the boundary between the Cretaceous and the Paleogene [~66 million years ago (Ma)], producing the ~200-km-wide Chicxulub impact crater (*1*–*4*). Impact markers at hundreds of sites globally co-occur with the deposition of the Cretaceous-Paleogene (K/Pg) boundary clay and include elevated abundances of siderophilic elements such as iridium, osmium, and nickel, as well as tektites and shocked quartz (*1*, *5*, *6*). The other disturbance was the eruption of an estimated ~500,000 km^3 of lava across much of India and into the deep sea in a large igneous province known as the Deccan Traps (*7*, *8*) during the K/Pg boundary–spanning magnetochron C29r [65.688 to 66.398 Ma, ~710,000 years long (*9*)]. Deccan volcanism was, like most flood basalt eruptions (*8*, *10*, *11*), episodic, with flows deposited in pulses throughout magnetochron C29r (*12*, *13*). That both volcanism and the impact event occurred within several hundred thousand years of the K/Pg extinctions is beyond reasonable doubt (*5*, *8*, *12*, *13*). However, many aspects of the mass extinction event are still uncertain, including the relative timing and magnitude of volcanic effects on the biosphere (*13*, *14*), the potential relationship between impact and volcanism (*7*, *13*, *15*), and whether impact or volcanism acted as the sole, primary, or joint drivers of extinction (*5*, *10*, *16*).

The case for the Chicxulub impact as a driver of K/Pg mass extinction includes processes hypothesized to operate during the days and decades after the collision. The bolide impact injected an estimated >50,000 km^3 of ejecta (*4*), as well as ~325 billion metric tons (Gt) of sulfur and ~425 Gt of CO_2 and other volatiles (*17*) into the atmosphere from the marine carbonate and anhydrite target rock of the Yucatan Peninsula (*5*, *18*). The combined effects of an expanding impact fireball and the reentry of molten ejecta from the skies (*19*) may have raised temperatures to the point of spontaneous combustion near the impactor and caused severe heat stress and even death many thousands of kilometers away from the impact site in minutes to days after impact (*20*). In the days to years that followed, nitrogen and sulfur vapors reacted to form nitric and sulfuric acids and, with CO_2 gases, acidified the oceans (*21*–*23*). Finally, models and empirical evidence suggest that the combination of dust and aerosols precipitated a severe impact winter in the decades after impact (*24*–*27*).

Notable though these environmental effects may be, some researchers question whether the Chicxulub impactor acted as the sole or main driver of the K/Pg mass extinction for three primary reasons. First, no single kill mechanism appears to explain the extinction patterns: Acidification (*28*, *29*) and primary productivity decline (*30*) [due to darkness and cold (*26*)] are favored in the marine realm, whereas heat exposure and loss of productivity [due to fires, darkness, and cold (*18*, *26*)] are favored in the terrestrial realm (*31*, *32*). Second, asteroid and comet impacts have occurred throughout the history of life [although likely none the size and force of Chicxulub (*33*) have taken place in the past ~500 million years (Myr)], but no other mass extinction is unambiguously linked to such a collision (*34*). Third, flood basalt volcanism is strongly implicated as the driver of two of the most destructive mass extinctions in the last ~half-billion years [the Permian-Triassic (P/T) and Triassic-Jurassic (T/J) extinctions], leading some to favor a similar role for Deccan volcanism in the K/Pg mass extinction (*35*). However, most episodes of flood basalt volcanism after the T/J extinction produced no increase in extinction rates (*36*), potentially owing to substantial Earth system changes that dampened the effects of flood basalts after the P/T extinction.

Questions regarding the role of Deccan volcanism in driving the K/Pg mass extinction arise because of the relative lack of evidence

[1]Department of Geology and Geophysics, Yale University, New Haven, CT 06511, USA. [2]Bundesanstalt für Geowissenschaften und Rohstoffe, 30655 Hannover, Germany. [3]GFZ German Research Centre for Geosciences, 14473 Potsdam, Germany. [4]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093, USA. [5]National Oceanography Centre Southampton, University of Southampton, Southampton SO14 3ZH, UK. [6]International Ocean Discovery Program, Texas A&M University, College Station, TX 77845, USA. [7]Departamento de Ciencias de la Tierra and Instituto Universitario de Ciencias Ambientales, Universidad Zaragoza, 50009 Zaragoza, Spain. [8]Géosciences Rennes, Université de Rennes 1, 35042 Rennes, France. [9]Department of Earth Sciences, University College London, London WC1E 6BT, UK. [10]Department of Geosciences, Pennsylvania State University, University Park, PA 16802, USA. [11]CEREGE, Université Aix-Marseille, 13545 Aix en Provence, France. [12]Institute for Rock Magnetism, University of Minnesota, Minneapolis, MN 55455, USA. [13]Institut für Planetologie, Universität Münster, 48149 Münster, Germany. [14]MARUM – Center for Marine Environmental Sciences, University of Bremen, 28359 Bremen, Germany. [15]Institute of Earth Sciences, Heidelberg University, 69120 Heidelberg, Germany. [16]Institut für Geophysik und Geologie, Universität Leipzig, 04103 Leipzig, Germany. [17]School of Geosciences, University of Edinburgh, Edinburgh EH8 9XP, UK. [18]Department of Geology & Geophysics, The University of Utah, Salt Lake City, UT 84112, USA. [19]Department of Biogeochemical Systems, Max Planck Institute for Biogeochemistry, 07745 Jena, Germany. [20]Department of Earth Sciences, Waseda University, Shinjyuku-ku, Tokyo 169-8050, Japan. [21]Department of Geosciences, Baylor University, Waco, TX 76798, USA. [22]Department of Earth Sciences, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA. [23]ConocoPhillips Company, Houston, TX 77079, USA. [24]Department of Geological Sciences and Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, CO 80309, USA. [25]School of Environment, Earth and Ecosystem Sciences, The Open University, Milton Keynes MK7 6AA, UK. [26]Harvard Society of Fellows, Harvard University, Cambridge, MA 02138, USA. [27]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA. [28]Department of Stratigraphy, Geological Survey of Denmark and Greenland (GEUS), DK-1350 Copenhagen K, Denmark. [29]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [30]Department of Earth and Environmental Sciences, Wesleyan University, Middletown, CT 06459, USA. [31]National Museum of Nature and Science, Tsukuba, 305-0005, Japan. [32]Department of Earth and Planetary Sciences, University of California, Santa Cruz, CA 95064, USA.
***Corresponding author. Email: pincelli.hull@yale.edu** †These authors contributed equally to this work.

for a volcanogenic driver. Despite advances in chronology, the timing of the most voluminous Deccan eruptions relative to the K/Pg extinctions remains unclear (*7*, *8*). In earlier studies, many researchers argued that most Deccan flood basalts (>85%) were emplaced in a relatively short interval before the K/Pg, starting around the C29r/C30n boundary (~66.39 Ma) and ending well before the K/Pg impact (*11*, *12*). In contrast, Renne *et al.* (*13*) and Sprain *et al.* (*8*) proposed that the vast majority of Deccan basalts were emplaced after the impact. Schoene *et al.* (*7*) largely agree with the basalt flow ages of Sprain *et al.* and Renne *et al.* (*8*, *13*) but place the K/Pg boundary higher in the lava pile (i.e., in the upper part of, or above, the Poladpur Formation) and therefore propose major pulses of emplacement immediately before and immediately after the impact (*7*).

Pre- and postimpact scenarios are debated in part because they are tied to different environmental disruption scenarios. Pre-event volcanism may have acted in concert with the impact to drive K/Pg extinctions (*10*), whereas post-event volcanism suggests a role for volcanism in the delayed recovery of biodiversity (*13*). For the environment and life, the main environmental effects of large igneous provinces are attributed to volatile release (*37*–*39*), not lava emplacement, and the magnitude of volcanic outgassing is not necessarily linked directly to the volume of erupted lava. If early eruptive phases of flood basalt volcanism have higher volatile concentrations, then most volatiles could have been released before the impact, even if most of the lava was emplaced afterward (*8*).

Here we provide constraints on Deccan Trap outgassing by comparing well-resolved and temporally detailed ocean drilling and global temperature records, with five modeled end-member scenarios for the timing, magnitude, and composition of outgassing (*40*). These comparisons allow us to consider the relative effects of Deccan Trap outgassing and bolide impact on the marine carbon cycle and biological change.

Marine environmental record of outgassing

Deccan Trap degassing released a mix of volatiles including SO_2, Cl and other halogens, and CO_2, with sulfur having perhaps the greatest direct effect on ecosystems through acidification and pronounced global cooling (>4.5°C) (*38*). However, the environmental effects of SO_2 would have been relatively short-lived (years to centuries at most) and difficult to detect in slowly accumulating deep-sea sediments. In contrast, the influence of CO_2 emissions should be clearly evident in marine sediments as a global warming event paired with a carbon isotope anomaly (*41*). We used this diagnostic fingerprint of CO_2 emissions as a proxy for the timing of potentially disruptive outgassing of sulfur (and other noxious gases) and to test which volcanic degassing scenarios are compatible with the observed record.

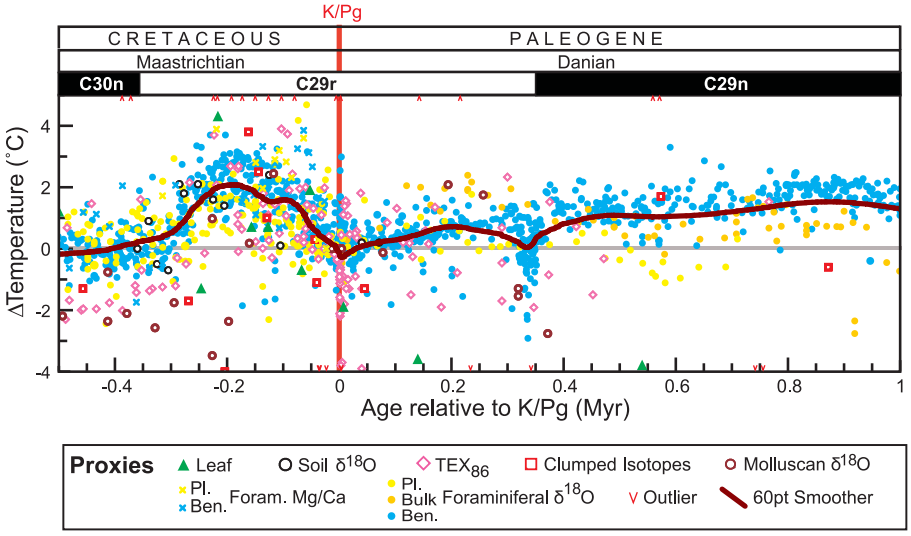


Fig. 1. Global temperature change across the K/Pg boundary. New and existing empirical temperature records from marine sediments (foraminiferal $\delta^{18}O$, foraminiferal Mg/Ca, and TEX_{86} measurements), shallow marine carbonates (clumped isotopes of mollusk carbonate), and terrestrial proxies (leaf margin analysis, biomarkers, clumped isotopes of mollusk carbonate) were aligned to a common age model (tables S2 and S3) and normalized to the latest Cretaceous temperature within each record. A 60-point fast Fourier transform (FFT) smoother of global temperature change is shown in dark red. Data are provided in tables S4 to S12. Some outlying data points do not fall within plot bounds but can be seen in figs. S1 to S16. Pl., planktonic; Ben., benthic; Foram., foraminiferal.

Two dominant features are clear in our global temperature compilation (Fig. 1) (*40*). First, marine and terrestrial records show a late Maastrichtian warming event of ~2°C, on average (figs. S1 to S16) (*42*–*44*), in the Cretaceous part of C29r that cooled back to pre-event temperatures before the K/Pg boundary (Fig. 1). Second, temperatures in the earliest Danian were comparable to those in the late Maastrichtian before the warming event, with temperatures gradually increasing to become >1°C warmer, on average, by ~600 thousand years (kyr) after the impact. Benthic foraminiferal oxygen isotope records, which typically track changes in global mean temperatures, show both of these features (Figs. 1 and 2 and fig. S13A), as do most other archives (figs. S1 to S16). The two exceptions, the bulk carbonate records and fish teeth phosphate records from El Kef (figs. S10C, S11, and S12), likely do not track global temperature for extinction-related reasons (*40*) and thus were excluded from our calculation of global mean temperatures.

Our multiproxy, astronomically tuned record from the North Atlantic site (*45*) has an especially complete Maastrichtian sequence and a millimeter-thick tektite layer at the K/Pg boundary (Fig. 2 and figs. S17 to S19). The record documents an excursion to lower $\delta^{13}C$ values in bulk sediments, coincident with $\delta^{18}O$ decline (a warming indicator) as well as a decline in osmium isotope values (Fig. 2 and figs. S20 and S21). Similar patterns are noted in records from the South Atlantic Walvis Ridge and the North Pacific Shatsky Rise (Fig. 2 and figs. S18 and S19) (*42*, *46*). The similarity of these records across three such widespread localities and four sites (Fig. 2) suggests that they provide a largely complete record of magnetochron C29r. Slight temporal offsets in the apparent onset and recovery from the latest Maastrichtian warming (among all sites) and in early Paleogene carbon isotope patterns at Shatsky Rise, due to short unconformities and/or the limitations of cyclostratigraphic age models, illustrate the current temporal uncertainties (Fig. 2). Temperature and atmospheric CO_2, as reflected in both our $\delta^{18}O$ and $\delta^{13}C$ anomalies and recent boron isotope records (*23*), returned to prewarming values in the very latest Maastrichtian. The most prominent feature in the records is the pronounced decline in $\delta^{13}C$ isotopes and change in sedimentary $CaCO_3$ content beginning at the K/Pg boundary (Fig. 2).

We investigated the timing of Deccan Trap outgassing by modeling the effects of CO_2 and sulfur emissions on long-term global temperatures using the geochemical box model LOSCAR (Long-term Ocean Sediment CArbon Reservoir v. 2.0.4) (*47*). Guided by published hypotheses for the timing and volume of trap emplacement, we tested five major Deccan Trap emission scenarios differing in the timing of volatile release: (i) case 1 (leading), with

the majority (87%) of degassing taking place before the K/Pg boundary [after (*10*)]; (ii) case 2 (50:50), with half of the degassing occurring before and half after the K/Pg boundary [after the lower estimate in (*8*)]; (iii) case 3 (punctuated), with four pulses including a major event just preceding the K/Pg boundary [after (*7*)]; (iv) case 4 (lagging), with the majority (87%) of degassing taking place after the K/Pg boundary [inverse case 1 pre- and post-outgassing volumes (*13*)]; and (v) case 5 (spanning), with emissions released evenly throughout magnetochron C29r [after (*12*)] (Table 1). All volcanic outgassing scenarios assume the same (i) initial climatic and oceanographic conditions [600 parts per million P_{CO_2} (partial pressure of CO_2) and climate sensitivity of 2° to 4°C per CO_2 doubling (*41*), LOSCAR's Paleogene ocean configuration and circulation, and marine [Mg^{2+}] of 42 mmol/kg and [Ca^{2+}] of 21 mmol/kg], (ii) K/Pg impact volatile release from the target rock (325 Gt S; 425 Gt CO_2) (*17*), (iii) upper and lower estimates for total volcanic outgassing volumes [4091 to 9545 Gt C and 3200 to 8500 Gt S (*10*) at constant ratios] (*40*), and (iv) extinction-related changes in the marine carbon cycle (*41*, *48*) (including reductions in both organic carbon and carbonate export and increases in intermediate-depth organic carbon remineralization; see Table 1) that taper back to pre-event values over 1.77 Myr after the extinction (*49*). In most outgassing scenarios, we assumed a common onset of Deccan degassing at the C30n/C29r boundary, following geochronology of the traps (*7*, *8*, *12*, *50*). In the age framework used to align the temperature records [i.e., GTS 2012 (*9*)], the C30n/C29r boundary is 358 kyr before the K/Pg boundary rather than the ~250 to 300 kyr indicated by the most recent $^{40}Ar/^{39}Ar$ and U-Pb geochronology (*7*, *50*). Simulations were initially tuned (*40*) to find the biological scenario (iv) that minimized mismatches between the data and model (figs. S22 to S27), and multiple scenarios for climate sensitivity and outgassing were considered in assessing goodness of fit (Figs. 3 and 4, figs. S25 and S28 to S32, and Table 2).

Three modeled scenarios differ distinctly from the observed pattern of temperature change (Fig. 3), and we thus consider them unlikely to represent the true outgassing history. Case 3 fails to reproduce the late Maastrichtian warming and shows a pronounced boundary-crossing warming event that is not supported by proxy data. In case 4, late Maastrichtian warming is too muted and early Paleocene warming is too pronounced, and in case 5 warming increases up to the K/Pg boundary, unlike in the empirical record (Fig. 3). Relatively poor model fit is also indicated by high mean

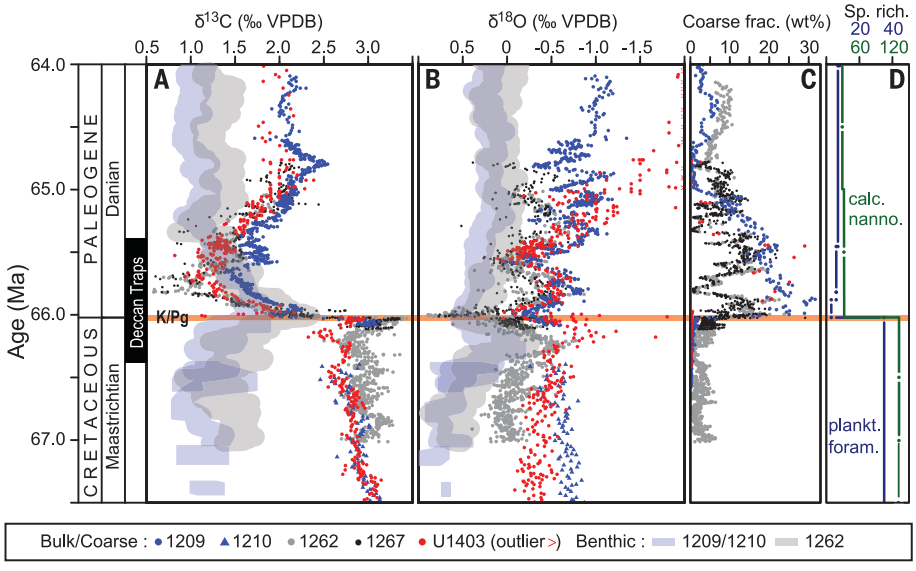


Fig. 2. K/Pg boundary dynamics at the best-resolved deep-sea sites globally: Shatsky Rise, Walvis Ridge, and J-Anomaly Ridge. High-resolution (**A**) carbon and (**B**) oxygen isotope dynamics in benthic foraminifera (transparent shaded areas) and bulk carbonate (discrete points) and (**C**) sediment composition (weight % coarse fraction) at Shatsky Rise (blue), Walvis Ridge (gray), and J-Anomaly Ridge (red). (**D**) Global records of nannofossil (green) and foraminifera [blue, from (*61*)] species richness (*40*). The major interval of Deccan Trap emplacement (estimated 93% of volume) is indicated at left by the black bar (*8*). Ocean drilling sites are listed by number. VPDB, Vienna Pee Dee belemnite; calc. nanno., calcareous nannofossil; plankt. foram., planktonic foraminifera; Sp. rich., species richness.

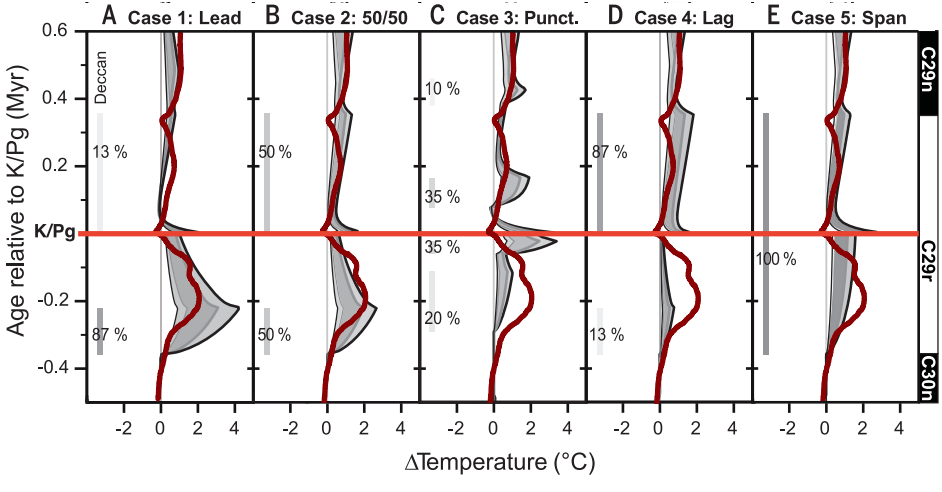


Fig. 3. Global temperature change across the K/Pg boundary compared to modeled temperature change in five scenarios for Deccan Trap outgassing. Outgassing scenarios include (**A**) case 1 (leading), with most outgassing before impact; (**B**) case 2 (50:50), with 50% outgassing before impact and 50% after impact; (**C**) case 3 (punctuated), with four pulses including a major event just before the K/Pg boundary; (**D**) case 4 (lagging), with most outgassing after impact; and (**E**) case 5 (spanning), with continuous outgassing throughout magnetochron C29r (Table 1). Each model scenario is represented by four lines (bounding a shaded region) delineating different combinations of climate sensitivity and volcanic outgassing: high degassing (9545 Gt C and 8500 Gt S) and 3°C per CO_2 doubling (thick gray line); high degassing and 4°C per doubling (thick black line); low degassing (4090 Gt C and 3200 Gt S) and 3°C per doubling (thin gray line); and low degassing and 2°C per doubling (thin black line). A 60-point FFT smoother of global temperature change (red line; see Fig. 1) is provided for comparison. The timing of Deccan outgassing assumed in each scenario is indicated by the bars at left in each panel, with the shading intensity of the bar denoting the proportion of outgassing in that interval.

absolute errors (MAEs) for cases 3 and 4 as compared with cases 1 and 2 (Table 2). The temporal dynamics of $\delta^{13}C$ in cases 3 and 5 also deviates from the empirical record (Fig. 4).

Only two outgassing scenarios produce modeled temperatures resembling those of the empirical records: the leading case (case 1) and the 50:50 case (case 2). We thus consider these the two most likely of the tested scenarios to represent Deccan Trap outgassing. In case 1, most CO_2 and SO_2 degassing occurred in the latest Maastrichtian, leading to global warming and subsequent cooling before the K/Pg. The relatively constant early Paleocene temperatures of case 1, with a gradual warming over the 600 kyr after the impact, are also consistent with empirical records (Figs. 1 to 3 and figs. S17 to S18). Case 2 (50:50) also matches the empirical temperature record well (Fig. 3), with the lowest MAEs of all cases (Table 2). The Late Cretaceous warming differs between case 1 and case 2 because of the reduced Late Cretaceous volcanic outgassing in the latter. Although uncertainty about climate sensitivity (*51*) and total Deccan Trap emissions (*10*, *12*) has a greater effect on modeled temperatures than the difference in outgassing volume between case 1 and case 2 (Fig. 3 and figs. S25 and S28), carbon isotopes also support case 2 as the more likely scenario (Fig. 4; see also MAEs in table S31).

The climatic effects of a major pulse (50%) of Deccan outgassing released over the ~350 kyr

Table 1. Model parameters for five focal Deccan outgassing scenarios tested in LOSCAR. Δ denotes change in value (reduction or increase). pre, before impact; post, after impact; Frac. int.-depth C_{org} remin., fraction of intermediate depth organic carbon remineralization.

	Case 1: Leading	Case 2: 50:50	Case 3: Punctuated	Case 4: Lagging	Case 5: Spanning
			Volcanic outgassing		
Pulse 1 (pre): Volume	87% of total high: 8305 Gt C, 7395 Gt S low: 3559 Gt C, 2784 Gt S	50% of total high: 4773 Gt C, 4250 Gt S low: 2045 Gt C, 1600 Gt S	20% of total high: 1909 Gt C, 1700 Gt S low: 818 Gt C, 640 Gt S	13% of total high: 1241 Gt C, 1105 Gt S low: 532 Gt C, 416 Gt S	100% of total high: 9545 Gt C, 8500 Gt S low: 4091 Gt C, 3200 Gt S
Timing	Starts: −358 kyr Ends: −218 kyr	Starts: −358 kyr Ends: −218 kyr	Starts: −290 kyr Ends: −110 kyr	Starts: −358 kyr Ends: −218 kyr	Starts: −358 kyr Ends: 355 kyr
Pulse 2 (pre): Volume			35% of total high: 3340 Gt C, 2975 Gt S low: 1431 Gt C, 1120 Gt S		
Timing			Starts: −60 kyr Ends: −20 kyr		
Pulse 1 (post): Volume	13% of total high: 1241 Gt C, 1105 Gt S low: 532 Gt C, 416 Gt S	50% of total high: 4773 Gt C, 4250 Gt S low: 2045 Gt C, 1600 Gt S	35% of total high: 3340 Gt C, 2975 Gt S low: 1431 Gt C, 1120 Gt S	87% of total high: 8305 Gt C, 7395 Gt S low: 3559 Gt C, 2784 Gt S	
Timing	Starts: 0 kyr Ends: 355 kyr	Starts: 0 kyr Ends: 355 kyr	Starts: 80 kyr Ends: 170 kyr	Starts: 0 kyr Ends: 355 kyr	
Pulse 2 (post): Volume			10% of total high: 955 Gt C, 850 Gt S low: 409 Gt C, 320 Gt S		
Timing			Starts: 390 kyr Ends: 430 kyr		
			Impact outgassing		
Volume	100% of total 115 Gt C, 325 Gt S	100% of total 115 Gt C, 325 Gt S	100% of total 115 Gt C, 325 Gt S	100% of total 115 Gt C, 325 Gt S	100% of total 115 Gt C, 325 Gt S
Timing	Starts: 0 kyr Ends: 1 kyr	Starts: 0 kyr Ends: 1 kyr	Starts: 0 kyr Ends: 1 kyr	Starts: 0 kyr Ends: 1 kyr	Starts: 0 kyr Ends: 1 kyr
			Biotic change		
Organic export flux Δ	50% reduction	50% reduction	50% reduction	50% reduction	50% reduction
$CaCO_3$ export flux Δ	42.5% reduction	42.5% reduction	42.5% reduction	42.5% reduction	42.5% reduction
Frac. int.-depth C_{org} remin. Δ	22% increase	22% increase	22% increase	22% increase	22% increase
Timing	Starts: 0 kyr immediately tapers Ends: 1770 kyr	Starts: 0 kyr immediately tapers Ends: 1770 kyr	Starts: 0 kyr immediately tapers Ends: 1770 kyr	Starts: 0 kyr immediately tapers Ends: 1770 kyr	Starts: 0 kyr immediately tapers Ends: 1770 kyr

immediately after the impact (case 2) were limited by extinction-related changes to the carbon cycle, including the reduction in $CaCO_3$ export from pelagic calcifiers to the seafloor. Marine $CaCO_3$ export indirectly affects atmospheric CO_2 by changing the distribution of carbon and alkalinity between the surface and the deep ocean and slows the removal of alkalinity from the system via $CaCO_3$ burial (*41*). The difference between cases 1 and 2 is almost imperceptible, with case 2 having slightly warmer (~0.25°C) early Danian temperatures than case 1. Notably, more-rapid Paleocene outgassing, such as that modeled in case 3 [after (*7*)], exceeds the capacity of the altered marine carbon cycle to absorb CO_2.

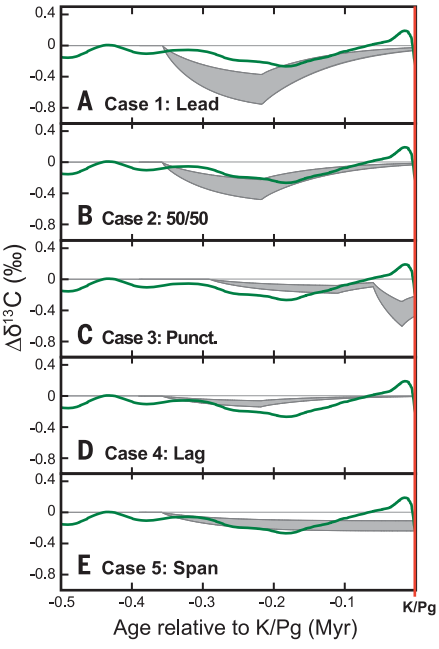


Fig. 4. Surface ocean $\delta^{13}C$ change across the late Maastrichtian warming compared to modeled $\delta^{13}C$ change in five scenarios for Deccan Trap outgassing. (**A** to **E**) Bulk carbonate $\Delta\delta^{13}C$ (20-point FFT smoother of data from Site U1403 and Site 1262) is shown against surface ocean $\delta^{13}C$ for end-member outgassing and climate sensitivity scenarios (gray shaded area) for each case, as detailed in Fig. 3. In each case, carbonate carbon isotopes are expressed as $\Delta\delta^{13}C$, relative to the late Maastrichtian high of 3.03‰ at 0.432 Myr before the onset of the CO_2 release (see also figs. S36 and S37).

Our results inform several boundary debates. First, if there was a large pulse of emplacement 20 to 60 kyr before the impact (*7*), then most CO_2 outgassing (and associated environmental impacts) must have preceded lava emplacement by several hundred thousand years. This would be before the eruption of the most voluminous stages of Deccan volcanism (i.e., before the Wai subgroup), as modeled for cases 1 and 2 [Figs. 3 and 4; see expanded discussion in (*40*)]. Second, roughly equal pre- and postimpact volcanic degassing is supported (case 2; Figs. 3 and 4), a hypothesized scenario in (*8*). However, our results are not consistent with most (>75%) volcanogenic degassing after impact [i.e., outgassing more similar to other eruptive volumes in (*8*, *13*)], because modeled warming is too muted in the Cretaceous and too pronounced in the early Paleocene (case 4) as compared with empirical records (Fig. 3). Third, impact-related volatile release from the target rock has a negligible climatic effect (fig. S24) and thus is unlikely to account for the pronounced warming in the first 100 kyr indicated by fish teeth $\delta^{18}O$ records (*52*). Instead, this record likely predominantly reflects changes in fish biology rather than temperature. Fourth, biotic recovery can account for the apparently gradual early Danian warming, as observed in marine records, if it begins at or shortly after impact and occurs over >1.5 Myr. This biotic recovery scenario reproduces the general pattern of change in $\delta^{13}C$ gradients (Fig. 2 and fig. S27), carbonate saturation state (Fig. 2C and fig. S27), and temperature but differs from recovery hypotheses that posit a delay in the onset of biological recovery for ~500 kyr or more (*40*, *49*, *53*).

No marine evidence for joint cause in mass extinction

The fossil record indicates no lasting, outsized, or cascading effect of the late Maastrichtian warming event on marine ecosystems of the sort that might predispose them to mass extinction by impact. First, we found no evidence for elevated extinction rates in the latest Cretaceous in marine taxa (table S1), excepting a contested record from Seymour Island, Antarctica (*54*, *55*). The scarcity of biostratigraphic datums in the Cretaceous portion of magnetochron C29r signifies a conspicuous lack of extinction in geographically widespread species, including planktonic foraminifera, nannoplankton, radiolarians, and ammonites (*9*). Second, late Cretaceous outgassing did not have a lasting effect on the community structure of well-fossilized taxa. Although range and community shifts coincided with warming, a shift back to the prewarming-like communities occurred before impact (table S1). Third, marine carbon cycle indicators ($\delta^{13}C$ and carbonate deposition) show no discernable effect of late Maastrichtian outgassing and warming on a major ecosystem function: the export and cycling of carbon. The $\delta^{13}C$ anomaly size [~0.2 to 0.3 per mil (‰); see also (*44*)] is consistent with a volcanogenic driver as in case 2 (Figs. 2 and 4 and fig. S28) given the magnitude of warming, without biological amplification.

In contrast, major and enduring changes to ecosystems coincided with the K/Pg impact. In deep-sea records, impact markers occur at the level of the abrupt mass extinction of >90% of planktonic foraminifera and 93% of nannoplankton species (Fig. 2). These groups exhibit rapid turnover and high dominance in community composition in the first 500 kyr of the Paleocene (*56*, *57*), when bulk carbonate $\delta^{18}O$ likely reflects community composition rather than surface ocean temperatures (Fig. 5 and

Table 2. Mean absolute error (MAE) and mean minimum absolute error (MMAE) of cases relative to the interpolated global temperature record. MMAE was calculated for each case by determining whether the empirical data fell outside of the temperature range bounded by the high- and low-outgassing scenarios, given a climate sensitivity of 3°C per CO_2 doubling, and, if so, by how much. MAEs were also calculated for each outgassing volume and climate sensitivity shown in Fig. 3. MMAEs and MAEs were calculated on a 20-kyr interpolated time step from 360 kyr before and 600 kyr after the K/Pg. Case 2 consistently has the lowest MAEs, and cases 1 and 2 have the lowest MMAEs. volc., volcanic outgassing; doub., doubling.

	MMAE	MAE (high volc., 3°C per CO_2 doub.)	MAE (high volc., 4°C per CO_2 doub.)	MAE (low volc., 3°C per CO_2 doub.)	MAE (low volc., 2°C per CO_2 doub.)
Case 1	0.25	0.46	0.65	0.50	0.58
Case 2	0.21	0.35	0.43	0.48	0.58
Case 3	0.45	0.59	0.65	0.58	0.64
Case 4	0.45	0.61	0.69	0.56	0.63
Case 5	0.29	0.40	0.44	0.53	0.61

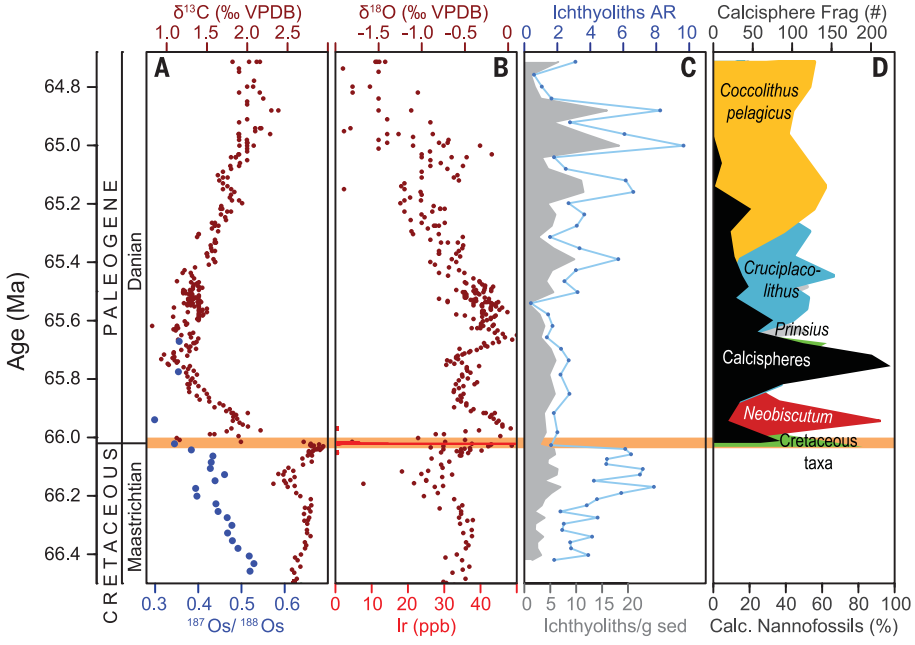


Fig. 5. Late Cretaceous warming and early Paleocene record of environmental and biotic change at IODP Site U1403, J-Anomaly Ridge, Newfoundland. A negative carbon isotope anomaly (**A**) coincides with late Cretaceous warming in $\delta^{18}O$ (**B**) and osmium isotope evidence for volcanism (A) at IODP Site U1403. The collapse in surface ocean $\delta^{13}C$ values (A) coincides with an iridium anomaly (B) and step change in fish tooth accumulation (**C**). The earliest Paleocene $\delta^{18}O$ values of bulk carbonate appear to be strongly influenced by vital effects driven by rapid turnover in the dominant calcareous nannofossil taxa (**D**) in sites globally (figs. S18, S34, and S35). Data are in tables S12, S16, S17, and S29. AR, accumulation rate; Frag, fragment; ppb, parts per billion; sed, sediment.

figs. S33 to S35). At the same time, tracers of the marine carbon cycle indicate a profound change in marine ecosystem function. The community structure of some groups, such as small fishes, which show no evidence of elevated extinction, changed permanently (*58*). The $\delta^{13}C$ composition of planktonic foraminifera and nannoplankton fell to or below that of benthic foraminifera at the iridium anomaly (Figs. 2 and 5 and figs. S34 and S35) (*43*, *49*). The loss or inversion of the $\delta^{13}C$ gradient typically maintained by the biological pump is unmatched in the fossil record of pelagic calcifiers (~170 Myr) and indicates that the K/Pg boundary impact had an outsized effect on the marine carbon cycle.

After the impact, an already-altered marine carbon cycle would have been needed to counteract the CO_2 emitted by a major postimpact pulse of outgassing, as in case 2 (Fig. 3), to avoid a warming event of the same magnitude as the Late Cretaceous warming event. This suggests that the major ecological change of the K/Pg mass extinction must have occurred before any major postimpact volcanism. Our modeling supports a scenario in which Deccan volcanism could have contributed to the aftermath of the impact and mass extinction, as in (*13*), if environmentally destructive gases such as SO_2, halogens, or sulfate aerosols contributed to (or drove) the persistence of unusual marine communities for the first ~500 kyr of the Paleocene. This might be particularly true if the evolution of the magma chamber led to higher sulfur content of later emissions, as in other eruption types (*59*). However, no observations document acidification coupled to extreme cold snaps in the earliest Paleocene, as predicted by this hypothesis, and there is no explanation for why SO_2 would have greater biotic effects in the well-buffered early Danian oceans than in the latest Maastrichtian oceans (figs. S1 to S18).

Outlook

We combined climatic, biotic, and carbon cycle records with modeled impact and outgassing scenarios and found support for a bolide impact as the primary driver of the end-Cretaceous mass extinction. Our analysis suggests that ~50% of Deccan Trap CO_2 outgassing occurred well before the impact, but it does not support the suggestion (*7*) that a large outgassing event took place a mere ~10 to 60 kyr before impact. This suggests a pronounced decoupling between CO_2 outgassing and lava flow emplacement, if the conclusions of Schoene *et al.* (*7*) are correct. Alternatively, our results support a relative impact and eruption chronology similar to the findings of Sprain *et al.* (*8*) and our best-supported, 50:50 outgassing scenario. The Late Cretaceous warming event attributed to Deccan degassing is of a comparable size to small warming events in the Paleocene and early Eocene that are not associated with elevated extinction or turnover (*43*, *60*), similar to what we find for the late Maastrichtian. We therefore conclude that impact and extinction created the initial opportunity for the rise of Cenozoic species and communities, but Deccan volcanism might have contributed to shaping them during the extinction aftermath.

REFERENCES AND NOTES

1. L. W. Alvarez, W. Alvarez, F. Asaro, H. V. Michel, *Science* **208**, 1095–1108 (1980).
2. A. R. Hildebrand *et al.*, *Geology* **19**, 867–871 (1991).
3. B. Collen *et al.*, *Biol. Lett.* **12**, 20150843 (2016).
4. J. Morgan *et al.*, *Nature* **390**, 472–476 (1997).
5. P. Schulte *et al.*, *Science* **327**, 1214–1218 (2010).
6. G. Ravizza, D. VonderHaar, *Paleoceanography* **27**, PA3219 (2012).
7. B. Schoene *et al.*, *Science* **363**, 862–866 (2019).
8. C. J. Sprain *et al.*, *Science* **363**, 866–870 (2019).
9. F. M. Gradstein, J. G. Ogg, M. D. Schmitz, G. M. Ogg, Eds., *The Geologic Time Scale 2012* (Elsevier, 2012).
10. A. L. Chenet *et al.*, *J. Geophys. Res.* **114**, B06103 (2009).
11. A. L. Chenet, X. Quidelleur, F. Fluteau, V. Courtillot, S. Bajpai, *Earth Planet. Sci. Lett.* **263**, 1–15 (2007).
12. B. Schoene *et al.*, *Science* **347**, 182–184 (2015).
13. P. R. Renne *et al.*, *Science* **350**, 76–78 (2015).
14. P. R. Renne *et al.*, *Science* **339**, 684–687 (2013).
15. M. A. Richards *et al.*, *Geol. Soc. Am. Bull.* **127**, 1507–1520 (2015).
16. E. Font *et al.*, *Earth Planet. Sci. Lett.* **484**, 53–66 (2018).
17. N. Artemieva, J. Morgan; Expedition 364 Science Party, *Geophys. Res. Lett.* **44**, 10180–10188 (2017).
18. S. P. S. Gulick *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19342–19351 (2019).
19. D. A. Kring, D. D. Durda, *J. Geophys. Res.* **107**, 5062 (2002).
20. J. Morgan, N. Artemieva, T. Goldin, *J. Geophys. Res. Biogeosci.* **118**, 1508–1520 (2013).
21. S. Ohno *et al.*, *Nat. Geosci.* **7**, 279–282 (2014).
22. T. Tyrrell, A. Merico, D. I. Armstrong McKay, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6556–6561 (2015).
23. M. J. Henehan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22500–22504 (2019).
24. J. Vellekoop *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7537–7541 (2014).
25. K. Kaiho *et al.*, *Sci. Rep.* **6**, 28427 (2016).
26. J. Brugger, G. Feulner, S. Petri, *Geophys. Res. Lett.* **44**, 419–427 (2017).
27. C. G. Bardeen, R. R. Garcia, O. B. Toon, A. J. Conley, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7415–E7424 (2017).
28. L. Alegret, E. Thomas, K. C. Lohmann, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 728–732 (2012).
29. B. J. Marshall, R. C. Thunell, M. J. Henehan, Y. Astor, K. E. Wejnert, *Paleoceanography* **28**, 363–376 (2013).
30. M. Aberhan, S. Weidemeyer, W. Kiessling, R. A. Scasso, F. A. Medina, *Geology* **35**, 227–230 (2007).
31. P. M. Sheehan, T. A. Hansen, *Geology* **14**, 868–870 (1986).
32. D. S. Robertson, M. C. McKenna, O. B. Toon, S. Hope, J. A. Lillegraven, *Geol. Soc. Am. Bull.* **116**, 760–768 (2004).
33. E. M. Shoemaker, *J. R. Astron. Soc. Can.* **92**, 297–309 (1998).
34. J. D. Archibald *et al.*, *Science* **328**, 973 (2010).
35. G. Keller, J. Punekar, P. Mateo, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **441**, 137–151 (2016).
36. S. V. Sobolev *et al.*, *Nature* **477**, 312–316 (2011).
37. M. T. Jones, D. A. Jerram, H. H. Svensen, C. Grove, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **441**, 4–21 (2016).
38. A. Schmidt *et al.*, *Nat. Geosci.* **9**, 77–82 (2016).
39. S. Self, S. Blake, K. Sharma, M. Widdowson, S. Sephton, *Science* **319**, 1654–1657 (2008).
40. Materials and methods are available as supplementary materials.
41. M. J. Henehan, P. M. Hull, D. E. Penman, J. W. B. Rae, D. N. Schmidt, *Philos. Trans. R. Soc. B* **371**, 20150510 (2016).
42. J. S. K. Barnet *et al.*, *Geology* **46**, 147–150 (2018).
43. J. S. K. Barnet *et al.*, *Paleoceanogr. Paleoclimatol.* **34**, 672–691 (2019).

44. L. Q. Li, G. Keller, *Geology* **26**, 995–998 (1998).
45. R. D. Norris, P. A. Wilson, P. Blum, Expedition 342 Scientists, in *Proc. IODP*, *342*, R. D. Norris, P. A. Wilson, P. Blum, Expedition 342 Scientists, Eds. (Integrated Ocean Drilling Program, 2014); http://publications.iodp.org/proceedings/342/342toc.htm.
46. N. Robinson, G. Ravizza, R. Coccioni, B. Peucker-Ehrenbrink, R. Norris, *Earth Planet. Sci. Lett.* **281**, 159–168 (2009).
47. R. E. Zeebe, *Geosci. Model Dev.* **5**, 149–166 (2012).
48. J. C. Zachos, M. A. Arthur, W. E. Dean, *Nature* **337**, 61–64 (1989).
49. H. S. Birch, H. K. Coxall, P. N. Pearson, D. Kroon, D. N. Schmidt, *Geology* **44**, 287–290 (2016).
50. C. J. Sprain, P. R. Renne, W. A. Clemens, G. P. Wilson, *Geol. Soc. Am. Bull.* **130**, 1615–1644 (2018).
51. E. J. Rohling *et al*., *Annu. Rev. Mar. Sci.* **10**, 261–288 (2018).
52. K. G. MacLeod, P. C. Quinton, J. Sepúlveda, M. H. Negra, *Science* **360**, 1467–1469 (2018).
53. S. D'Hondt, P. Donaghay, J. C. Zachos, D. Luttenberg, M. Lindinger, *Science* **282**, 276–279 (1998).
54. J. D. Witts *et al*., *Nat. Commun.* **7**, 11738 (2016).
55. T. S. Tobin, *Sci. Rep.* **7**, 16317 (2017).
56. P. M. Hull, R. D. Norris, T. J. Bralower, J. D. Schueth, *Nat. Geosci.* **4**, 856–860 (2011).
57. J. J. Pospichal, in *The Cretaceous-Tertiary Event and Other Catastrophes in Earth History*, G. Ryder, D. E. Fastovsky, S. Gartner, Eds. (GSA Special Paper 307, Geological Society of America, 1996), pp. 335–360.
58. E. C. Sibert, M. Friedman, P. M. Hull, G. Hunt, R. D. Norris, *Proc. R. Soc. B* **285**, 20181194 (2018).
59. M. Edmonds, *Philos. Trans. R. Soc. London Ser. A* **366**, 4559–4579 (2008).
60. P. F. Sexton *et al*., *Nature* **471**, 349–352 (2011).
61. R. D. Norris, in *Palaeobiology II*, D. E. G. Briggs, P. G. Crowther, Eds. (Blackwell Science, 2001), pp. 229–231.

ACKNOWLEDGMENTS

This research used samples and/or data provided by the International Ocean Discovery Program (IODP), which was sponsored by the U.S. National Science Foundation and participating countries under management of Joint Oceanographic Institutions, Inc., and its predecessors—the (Integrated) Ocean Drilling Program and the Deep Sea Drilling Program. We thank the *JOIDES Resolution* crew of IODP Expedition 342 and W. Hale and A. Wuelbers for help with sampling. We also thank the many centers and staff scientists who enabled the measurements, including L. Elder in the Hull laboratory (Yale University), B. Erkkila and M. Wint at the Yale Analytical and Stable Isotope Center, D. Andreasen at the UCSC Stable Isotope Laboratory, and F. Demory (CEREGE) for help with magnetic data production and processing. This work benefited from helpful discussions with J. Dinarès-Turell, the insights of C. B. Keller, and the comments of four anonymous reviewers. **Funding:** IODP USSSP Post-Expedition Activity award and Yale University support to P.M.H.; Deutsche Forschungsgemeinschaft (DFG) funding (grants BO2505/8-1 and EH 89/20-2) to A.B.; Yale Peabody Museum support to M.J.H.; Spanish Ministry of Economy and Competitiveness and FEDER funds (CGL2017-84693-R) to L.A.; DFG funding (grant VO687/14) to S.J.B.; a Richard Foster Flint Postdoctoral Fellowship (Department of Geology and Geophysics, Yale University) to D.E.P.; DFG funding (grant number FR2544/2) to O.F.; NSF funding (EAR-132552) and American Chemical Society Petroleum Research Fund (grant PRF#52822-DN18) to D.J.P.; DFG funding (grants RO1113/3, RO1113/4, and RO1113/8) to U.R.; the NASA Exobiology Program (grant NNX09AM88G) to R.E.S.; the Danish Council for Independent Research/Natural Sciences (DFF/FNU) (grant 11-107497) to K.K.Ś.; NSF funding (OCE 1536611) to E.T.; DFG funding (grant WE5479/3) to T.W.; and a NERC (NE/K006800/1) and Royal Society Wolfson award to P.A.W. **Author contributions:** P.M.H. conceived and co-led the study, drafted the manuscript, contributed to model design, generated empirical data, and edited data tables and figures. A.B. co-led the study; coordinated data generation, reporting, figures, and tables; generated empirical data; and substantially contributed to study design and text. D.E.P. led LOSCAR modeling and substantially contributed to study design and text. M.J.H. compiled and aligned age models for the global temperature compilation, prepared related tables and figures, and substantially contributed to study design and text. R.D.N., P.A.W, and P.B. led IODP Expedition 342, with R.D.N. and P.A.W. substantially contributing to study design and text. L.A., S.J.B., P.R.B., T.J.B., C.C., A.D., B.D., O.F., S.J., H.K., D.K., P.C.L., D.L., I.M., K.M., D.J.P., G.E.R., U.R., J.S., J.D.S., E.C.S., K.K.Ś., R.E.S., E.T., T.W., J.H.W., and T.Y. contributed empirical datasets, figures, and related analyses, interpretations, and text. L.A., P.R.B., T.J.B., O.F., D.K., P.F.S., J.S., E.T., T.W., J.H.W., and J.C.Z. substantially contributed to ideas and/or text. All authors read and approved the final text. The primary contribution of R.D.N., P.A.W., P.B., S.J.B., P.R.B., C.C., A.D., O.F., S.J., H.K., P.C.L., D.L., I.M., K.M., G.E.R., U.R., J.S., P.F.S., E.C.S., K.K.Ś., R.E.S., T.W., J.H.W., and T.Y. was the IODP Expedition 342 K/Pg boundary investigation. In the byline, the first seven authors appear in order of contribution; all others are listed alphabetically.
Competing interests: The authors have no competing interests.
Data and materials availability: All data are available in the manuscript and the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/266/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S44
Tables S1 to S31
References

24 June 2019; accepted 5 December 2019
10.1126/science.aay5055

PALEONTOLOGY

A high-resolution summary of Cambrian to Early Triassic marine invertebrate biodiversity

Jun-xuan Fan[1,2], Shu-zhong Shen[1,2,3]*, Douglas H. Erwin[4,5], Peter M. Sadler[6], Norman MacLeod[1], Qiu-ming Cheng[7], Xu-dong Hou[1], Jiao Yang[1], Xiang-dong Wang[1], Yue Wang[2], Hua Zhang[2], Xu Chen[2], Guo-xiang Li[2], Yi-chun Zhang[2], Yu-kun Shi[1], Dong-xun Yuan[2], Qing Chen[2], Lin-na Zhang[2], Chao Li[2], Ying-ying Zhao[1]

One great challenge in understanding the history of life is resolving the influence of environmental change on biodiversity. Simulated annealing and genetic algorithms were used to synthesize data from 11,000 marine fossil species, collected from more than 3000 stratigraphic sections, to generate a new Cambrian to Triassic biodiversity curve with an imputed temporal resolution of 26 ± 14.9 thousand years. This increased resolution clarifies the timing of known diversification and extinction events. Comparative analysis suggests that partial pressure of carbon dioxide (P_{CO_2}) is the only environmental factor that seems to display a secular pattern similar to that of biodiversity, but this similarity was not confirmed when autocorrelation within that time series was analyzed by detrending. These results demonstrate that fossil data can provide the temporal and taxonomic resolutions necessary to test (paleo)biological hypotheses at a level of detail approaching those of long-term ecological analyses.

Understanding patterns of global diversity can reveal the history of the biosphere and relations between environmental changes and diversity fluctuations, and can provide insights into how the fossil record might inform current biodiversity concerns. Early global-scale quantitative analysis identified what have come to be known as the "big five" mass extinctions. However, such efforts depend on the quality and temporal resolution of paleontological data, which have improved substantially since the 1990s, most recently through the intensive data compilation of the Paleobiology Database (*1*–*6*). Analyses of those data have increased our understanding of paleobiodiversity (*4*, *7*–*10*).

Previous deep-time paleobiodiversity reconstructions (*1*, *11*) were limited by coarse age determinations of taxon occurrences. The relatively long and uneven duration of age bins (stage or series level) used in these studies imposed complexly structured limits on resolving power across different intervals. Resolutions were generally no better than 8 to 11 million years (Myr) with standard deviations of 2.4 to 3.2 Myr, although some trials have been made to achieve better resolution for the early Paleozoic (*6*). Taxon age assignments were subject to error, not equally applicable to all clades, and quickly became outdated by new correlations or updated age estimates. Previous analyses have also been performed at taxonomically broad and phylogenetically suspect family or genus levels. Such resolutions are often too crude and imprecise to assess diversification rates or patterns associated with various global events (gradual, stepwise, or abrupt) and may mask multiple events as well as finer-scale fluctuations (*7*, *12*, *13*).

Here, we used a new parallel computing implementation of the constrained optimization method (CONOP.SAGA) run on the Tianhe II supercomputer. This approach uses inferred stratigraphic correlations to construct composite biodiversity curves for Cambrian to Triassic marine invertebrate genera and species (Fig. 1) and has demonstrated the capacity to establish finely resolved, traceable time zones over wide geographic areas (*14*).

Data and methods

Data compilation and standardization were conducted through the Geobiodiversity Database (*15*). This database is particularly suitable for biodiversity studies because, unlike the Paleobiology Database (Fig. 2), it is based on section data and provides quality control at a bed-by-bed level using an online, interactive system for recording expert taxonomic opinions (*15*). Taxonomic and age assignments used in this investigation were vetted by a team of 11 paleontologists, who checked and updated each taxonomic record. We also cross-checked these species names for synonyms.

Because the Geobiodiversity Database records local taxon occurrences and their positions in stratigraphic sections, we were able to construct a composite sequence of assemblages and calibrate this sequence to a current estimate of the geological time scale using the best available chronostratigraphic data (*16*). Our study focused on marine invertebrates and used data from 3766 published stratigraphic sections, including 266,110 local records of the stratigraphic ranges of 45,318 taxonomic units, covering all Chinese Cambrian to Lower Triassic tectonic blocks (fig. S1). Although our data were largely derived from Chinese sections, the tectonic blocks on which they reside were situated in paleolatitudes stretching from southern Gondwanan to northern Boreal realms (*17*). Accordingly, these data reflect global biodiversity patterns (figs. S2 and S3).

Our initial analyses revealed that the rarity of Silurian–Devonian data in China (due to worldwide regression) hampered regional and global correlations. Consequently, we added a small amount of European Silurian–Devonian data to improve the correlations in this interval. These additional data did not alter the generality of our results because different Chinese tectonic blocks were located in different regions during the Paleozoic, with some residing close to Europe (fig. S3). Our study interval terminated at the late Middle Triassic marine regression.

Taxonomic names in open nomenclature, questionable taxa, and taxa unidentifiable to the species level were not included. Species recorded from only one locality were also removed to avoid the "monograph effect" (*18*). The resulting final dataset contained 116,060 local records of total stratigraphic ranges of 11,268 species from 3112 published stratigraphic sections.

To avoid the need to use coarse time bins, we used constrained optimization (CONOP) (*19*) stratigraphic correlation to reconstruct the Paleozoic biodiversity history of marine invertebrates. The CONOP correlation method, which applies a simulated annealing algorithm to infer a globally optimized sequence of stratigraphic datums, has been used previously for local high-resolution biochronostratigraphic studies (*14*, *20*, *21*). However, the original CONOP algorithm (*22*, *23*) did not support parallel or high-performance computing and it would have required dozens of years to calculate one CONOP composite for this dataset. To overcome this "big data" problem, we modified the original CONOP algorithms to parallelize the sequencing problem. We also designed a special hybrid strategy of simulated annealing and genetic algorithm for the parallel computing application, CONOP.SAGA (*16*).

CONOP.SAGA iteratively compares species ranges from many local range charts to assemble the global first and last occurrence datums into a single, global, best-fit sequence, thereby reducing the effect of local-section

[1]School of Earth Sciences and Engineering, Nanjing University, Nanjing 210023, China. [2]LPS, Nanjing Institute of Geology and Palaeontology and Center for Excellence in Life and Paleoenvironment, Chinese Academy of Sciences, Nanjing 210008, China. [3]Key Laboratory of Continental Collision and Plateau Uplift, Institute of Tibetan Plateau Research and Center for Excellence in Tibetan Plateau Earth Sciences, Chinese Academy of Sciences, Beijing 100101, China. [4]Department of Paleobiology, National Museum of Natural History, Washington, DC 20013, USA. [5]Santa Fe Institute, Santa Fe, NM 87501, USA. [6]Department of Earth Sciences, University of California, Riverside, CA 92521, USA. [7]School of the Earth Sciences and Resources, China University of Geosciences, Beijing 100083, China.
*Corresponding author. Email: szshen@nju.edu.cn

incompleteness substantially (*22*). As is often the case in complex optimal modeling problems, CONOP.SAGA does not yield a single solution, but rather a set of equally optimal solutions. To obtain robust results, we created a medium-sized dataset and made >10 complete CONOP.SAGA calculations. If the results were similar for each run, then this suggested that the true global optimal solution had been found. Basic settings of the CONOP.SAGA parameters for the full dataset were estimated from these results. We then uploaded the full dataset and ran it while varying these parameters. Each calculation took 20 to 60 hours to complete and produced one composite sequence. Structured increase in these parameters allowed us to determine the effect of their variation on the diversity curve. Despite >1.2 billion iterations per run, all runs produced consistent results (>98.6%) with nearly indistinguishable variations. We then used the best (i.e., lowest misfit) (*19*) of the final three complete-dataset solutions to construct the final diversity curve estimate (Fig. 1).

Imprecision in estimating age and duration of biozones or other time intervals was prevented because local, isotopically dated species' first and last appearance levels in the composite sequence were calibrated directly by regression against the high-precision geochronologic dates used to calibrate the Cambrian to Early Triassic geologic time scale (*24*, *25*) with estimated ages for the bases of major international standard biozones (fig. S5 and table S2) (*24*). This calibrated composite is composed of 11,326 discrete temporal levels for the interval from 538.85 to 244.41 million years ago (Ma), yielding a mean temporal resolution of ~26.0 ± 14.9 thousand years (kyr).

Both species- and genus-level diversity summaries were assembled using an unbinned method (*16*, *20*) to avoid issues associated with different counting strategies (*26*). A simple diversity curve was created by counting the number of species and genera occurring at each temporal level (Fig. 1A). Fossil data are inevitably biased by incomplete preservation and sampling across time, geographic range, and environmental setting. We used a bootstrap technique to estimate the range of variation of the species-diversity curve (Fig. 1A) (*20*). The diversity at each point was also standardized by average diversity and the number of sections to reduce effects of sampling effort (figs. S6 and S7) (*16*).

Results

Our results (Fig. 1A and fig. S6) revealed a sharp increase in diversity associated with the Cambrian explosion, a pause through the late Cambrian Steptoean positive carbon isotope excursion event (*27*), followed by a nearly threefold increase in species diversity during the Early Ordovician. Species/genus ratio data suggest that the Great Ordovician Biodiversification Event (GOBE) reflects species-level diversification, which resulted in a facultative expansion of marine ecosystems. By contrast, the Cambrian radiation was associated mainly with increases at the genus level or among higher taxa (Fig. 1A). The GOBE (*6*, *28*) is evident as a sustained 29.72-Myr-long diversity increase (from 497.05 to 467.33 Ma) until the Middle Ordovician. By contrast, the Alroy *et al.* diversity curve [(*3*); Fig. 1] showed a steady radiation from the early Cambrian until the Emsian (Early Devonian) with a minor drop across the Ordovician–Silurian transition (Fig. 2A).

The end-Ordovician mass extinction is evident as a rapid diversity decline from the late Katian to the Hirnantian; two phases corresponding to the development and disappearance of the *Hirnantia* fauna were previously identified (*18*, *29*), but that distinction is less evident in our Chinese species-level summary (Fig. 1A and fig. S8). This interval may represent either a single event (*30*) or a community-level turnover, presumably in response to Hirnantian glaciation. An immediate recovery and radiation occurred near the Ordovician–Silurian boundary (444.66 Ma) and persisted until the early Silurian (Telychian: 437.08 Ma). Previous studies have not recognized this earliest Silurian radiation (*4*).

The Late Devonian Frasnian–Famennian extinction was broadly evident in the Alroy *et al.* curve (*4*) with a rebound in the Famennian.

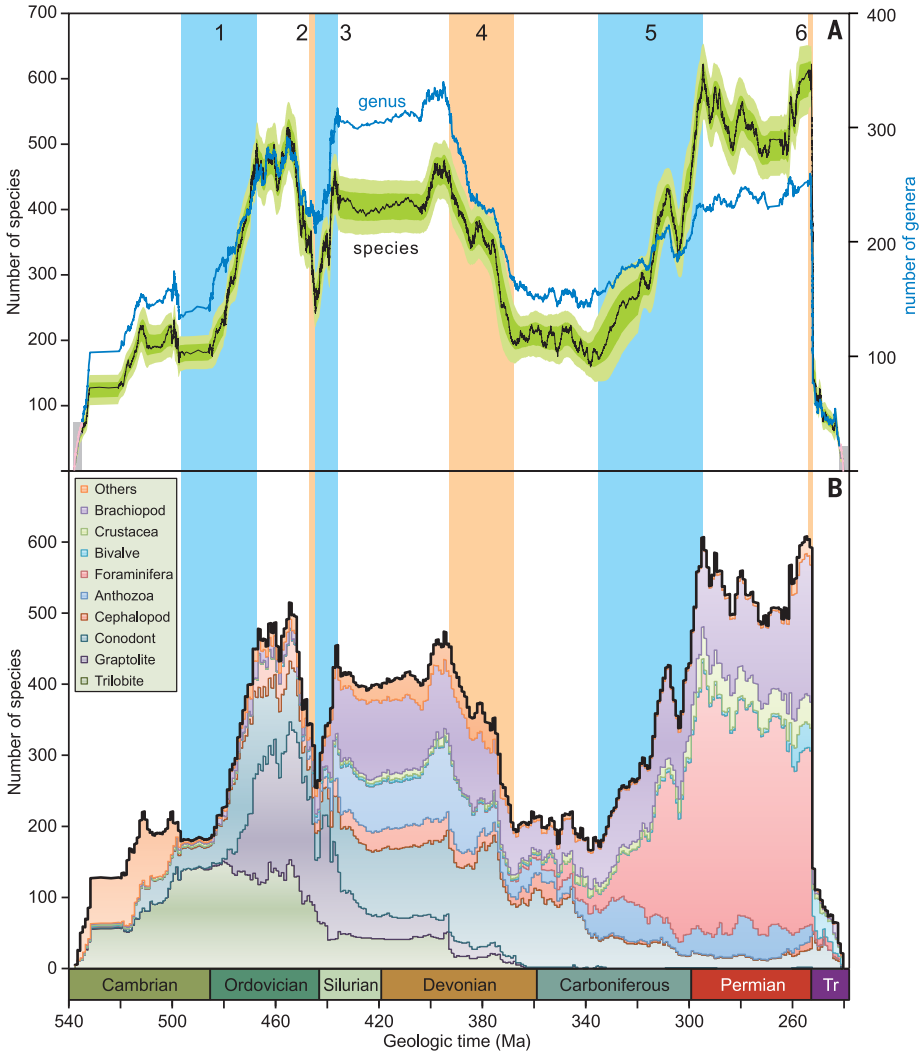


Fig. 1. General trajectories of Paleozoic genus and species diversity and species diversity for 10 major fossil groups. (**A**) Genus and species diversity. (**B**) Species diversity. The light and dark green shading in (A) represent 1σ and 2σ standard deviations, respectively, which are based on 500 bootstrap runs. 2σ approximately equals the 95% confidence interval. Gray bars on either side show the buffer zones with edge effects. 1, GOBE; 2, end-Ordovican mass extinction; 3, early Silurian radiation; 4, Middle to Late Devonian diversity decline; 5, late Carboniferous–early Permian biodiversification event; 6, end-Permian mass extinction.

By contrast, our higher-resolution summary shows the Frasnian–Famennian event to be embedded within a protracted diversity decline that began in the Eifelian (392.72 Ma) and continued to ~368.24 Ma with no subsequent major recovery. No discrete Frasnian–Famennian event (*31*, *32*) was evident in our results (Fig. 1A and fig. S6).

The Late Paleozoic ice age was accompanied by a previously unrecognized great biodiversification event (the Carboniferous–Permian Biodiversification Event) comparable to the GOBE on the basis of both the increase of species richness and species/genus ratio (Fig. 1A and fig. S9). The Alroy *et al*. result has this biodiversification event extending from the early Cisuralian into the middle Guadalupian (Figs. 1A and 2A and fig. S8) (*4*).

We found little evidence for an end-Guadalupian mass extinction, confirming previous qualitative results (*33*). The temporal resolution of earlier global compilations was too coarse to reveal the abrupt nature of the end-Permian mass extinction, but this is resolved clearly in our results. Both species and genus diversities declined rapidly from 252.73 to 251.95 Ma, followed by a sudden drop over ~63 kyr (Fig. 1A), which is consistent with the duration for this event calibrated with high-precision geochronology based on the Meishan section (table S1 and fig. S6) (*34*). This extinction was followed immediately by a minor Late Early Triassic radiation at 249.58 Ma, but marine diversity remained low through the Early and Early Middle Triassic (fig. S6).

To compare these results with diversity curves from the Sepkoski and Paleobiology databases, we generated generic and species-diversity patterns using ~10-Myr time bins with durations similar to those used by Alroy *et al*. (*4*). This coarse graining of the data yielded generally comparable patterns for the GOBE, the Early Silurian Radiation, and the Carboniferous–Permian Biodiversification Event, along with three major diversity crises (Late Ordovician, Middle to Late Devonian, and end-Permian) (Fig. 2). More importantly, this comparison shows the dependence of paleobiodiversity estimation on temporal resolution. If our high-resolution diversity pattern (Fig. 1A and fig. S6) is compared with the patterns of the same data using the ~10 Myr-bins, binning spreads out the end-Ordovician mass extinction and the Early Triassic radiation largely disappears (Fig. 2). These differences reflect the effects of low temporal resolution and uneven time-bin durations, varying from 4.97 to 32 Myr (fig. S11).

We used different time-bin durations (from 0.1 to 10 Myr) to calculate a series of diversity curves (Fig. 3) in addition to parsing our data into the unequal time bins used by Alroy *et al*. (*4*). The diversity pattern produced using 10-Myr bins is very close to these (*4*) but the influence of coarse and uneven sampling remained evident. Bins of up to 0.5-Myr duration retained a largely distortion-free representation of fine-scale structure in the present paleobiodiversity analysis (~300 Myr; Fig. 3). Previous studies suggest that resolutions as high as 0.1 to 0.5 Myr are needed to test causal-association hypotheses because this is the estimated duration of many physical perturbations to environmental states (e.g., large igneous province eruptions, bolide impacts, ocean anoxia events).

In terms of Sepkoski's three marine evolutionary faunas (*35*) (fig. S8) and different fossil groups (Fig. 1B and fig. S10), the Cambrian fauna represent the bulk of Cambrian biodiversity and remained substantial during the Ordovician, although they were later overwhelmed by diversification of the Paleozoic fauna. Trilobites, graptolites, and hyolithids lost their dominance after the Silurian.

The Paleozoic evolutionary fauna, consisting largely of brachiopods and other epifaunal filter feeders, dominated marine habitats from the Ordovician through the Devonian diversity decline. Conodont species diversity reveals two distinct radiations, one in the Early and Middle Ordovician and the other from the Middle Devonian to the late Carboniferous. Conodonts declined gradually during the early Carboniferous and experienced considerable losses during the Late Paleozoic Ice Age. Brachiopods experienced a rapid latest Ordovician to early Silurian diversification. This radiation was interrupted by a minor fluctuation in response to waxing and waning of the Hirnantian unipolar ice sheet. Brachiopod diversity declined during the Middle and Late Devonian before a steady increase began from the late early Carboniferous. The Lopingian had the most diverse brachiopod faunas of the Paleozoic, but this clade experienced catastrophic losses during the end-Permian mass extinction (*36*). Fusulinids exhibit a substantial diversification from the mid-Carboniferous to early Cisuralian and constitute an important component of the Carboniferous–Permian Biodiversification

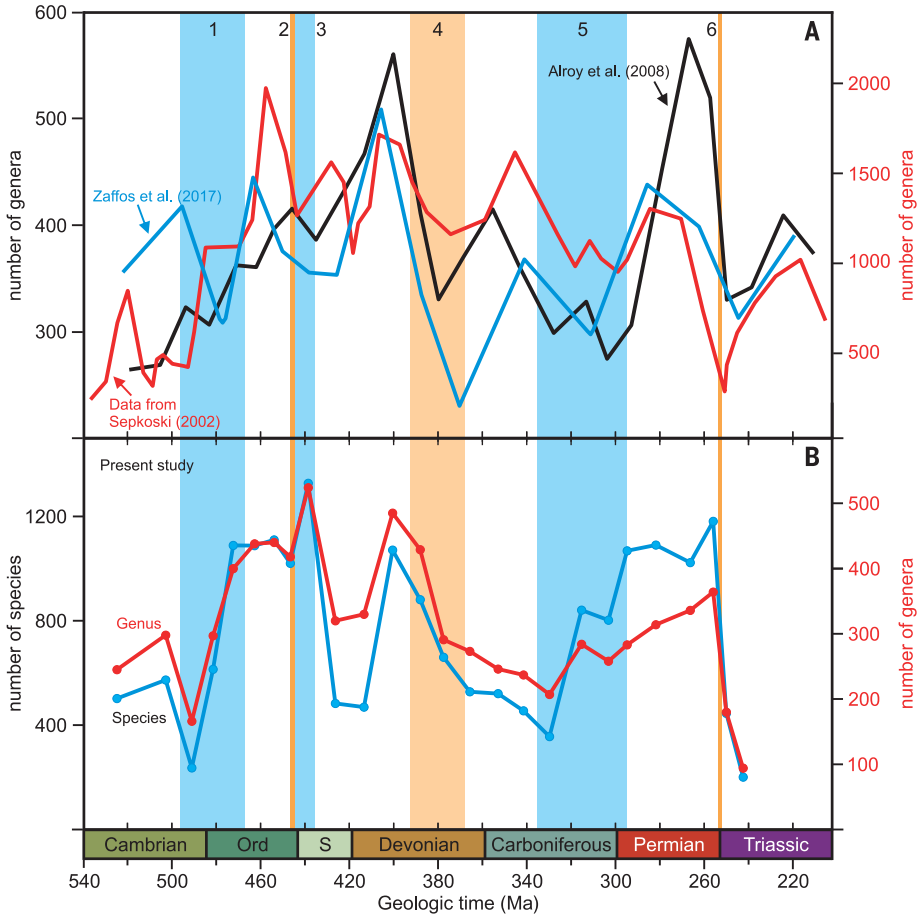


Fig. 2. Comparisons between previous diversity curves and the present study for the Paleozoic. (**A**) Diversity curves of previous studies. (**B**) ~10-Myr-binned species and genus diversity curves based on the present dataset. The high-resolution unbinned species diversity in Fig. 1 has been rescaled to the 26 uneven time bins adopted by Alroy *et al*. (*25*). The two *y*-axes on the right only refer to the red curves.

Event along with brachiopods and crustaceans. Fusulinid diversity declined slightly until the Lopingian, with the clade disappearing during the end-Permian mass extinction (Fig. 1B and fig. S10).

Mechanisms of diversity change

Many factors have been invoked as potential driving factors of biodiversity changes, including changes in paleoclimate, sea level, nutrient flux, ocean-atmospheric circulation, total habitable area, and intercontinental connectivity (*8*, *10*, *37*–*39*). There are two challenges to establishing the relationships between environmental drivers and diversity: the need for high-resolution proxy data for appropriate environmental indicators (Fig. 4) and the confounding problems of potential cross-correlation of autocorrelated time series. Data of appropriate resolution are mostly lacking, particularly for changes in global and regional climate (*38*, *40*). Currently available data suggest that the relationship between diversity and climate during the Paleozoic was complex. The GOBE and the Carboniferous–Permian Biodiversification Event were associated with climatic cooling (*41*, *42*), which stands in contrast to the modern association between warm climates and diversity (Fig. 4). The early Silurian radiation was associated with overall warming until the middle Llandovery (*43*). The late Asselian (294.19 Ma) diversity peak coincided with the acme of Late Paleozoic Ice Age (*41*, *44*). Species diversity declined from the Sakmarian to middle Guadalupian as climate ameliorated (*45*). The end-Ordovician extinction was associated with a short glaciation (*46*, *47*).

The Middle to Late Devonian diversity decline also had a complex temperature history. Paleotemperature increased from the Eifelian–Givetian transition to the Frasnian–Famennian boundary, then decreased after the Frasnian–Famennian boundary (*48*). On the basis of the latest $\delta^{18}O_{apatite}$ data, the Lower and Upper Kellwasser events were each associated with distinct cooling events (*49*). The end-Permian mass extinction was followed immediately by a rapid warming of 8 to 10°C (*50*, *51*), but low diversity in the Early Triassic coincided with a lethal "hothouse" (*52*).

The carbon isotope record reflects changes in diversity and abundance that affect the global carbon cycle (*53*). Therefore, large-scale biotic events are often associated with large carbon isotope excursions (*54*). The end-Ordovician mass extinction is associated with an ~4‰ positive excursion of $\delta^{13}C_{carb}$ (*55*), and the end-Permian mass extinction includes an ~5‰ negative excursion (*34*). Carbon isotope excursions have been reported from the Late Devonian Frasnian–Famennian event (*56*) but do not appear to coincide with distinct diversity changes (*54*).

Changes in the atmospheric concentrations of P_{CO_2} have had a major impact on earth system dynamics, but there have been major discrepancies between previous reconstructions of secular trends (*57*). A recent long-term P_{CO_2} reconstruction (*58*) and the stable carbon isotopic fractionation associated with photosynthesis (*59*) show a secular trend similar to that documented by our late Silurian to Early Triassic diversity pattern (Fig. 4, F and G). The increasing trend in P_{CO_2} is associated

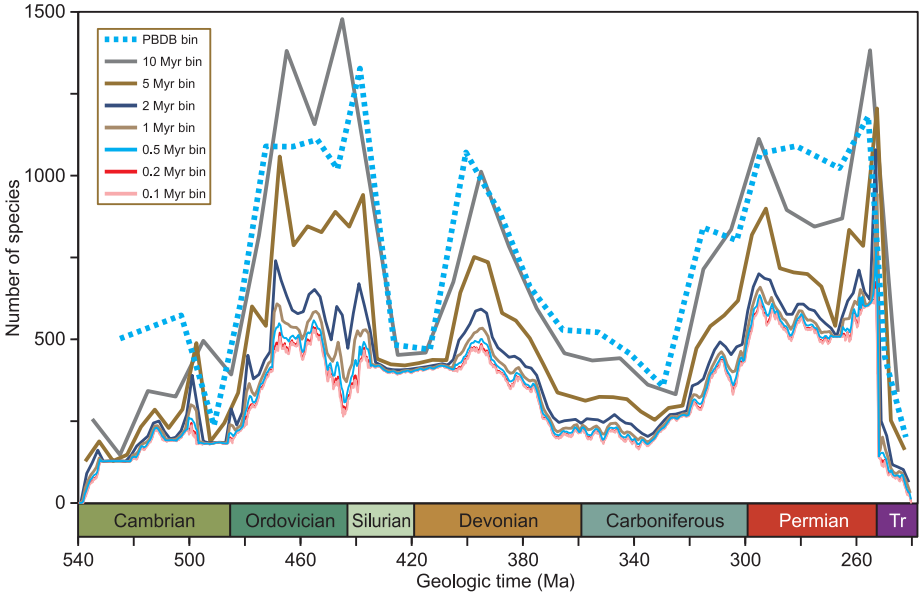


Fig. 3. Composite diversity patterns produced in different even time bins (0.1 to 10 Myr) to show the biases and effects of different temporal resolutions on paleobiodiversity estimation. Dashed blue line represents the result fitting in the unequal time bins adopted by Alroy *et al.* (*4*).

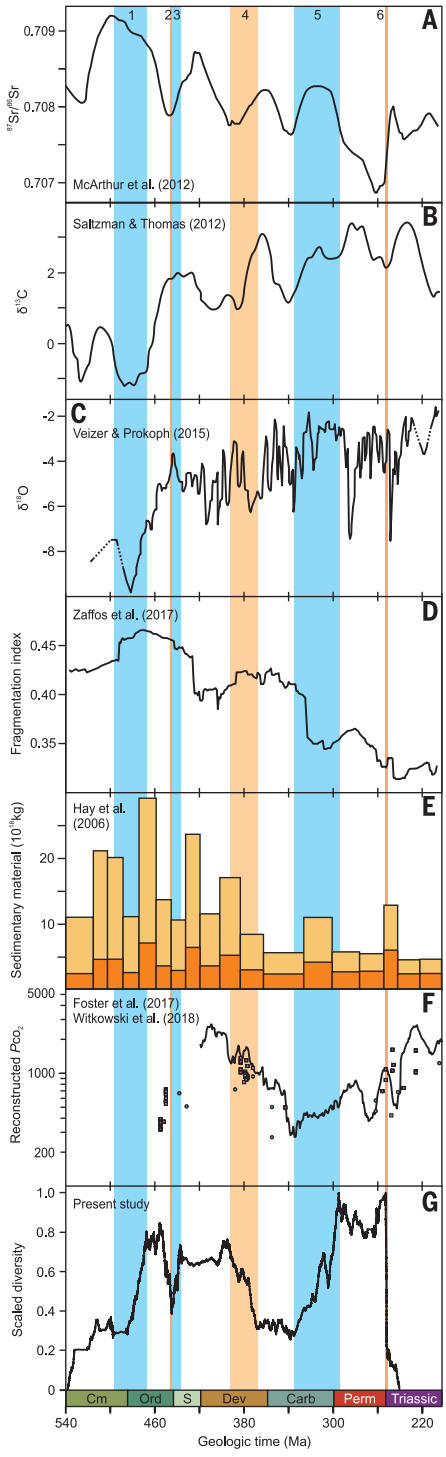


Fig. 4. Correlation between the species-diversity trajectory and the trends of multiple environmental proxies. (**A**) $^{87}Sr/^{86}Sr$ ratio (*60*). (**B**) $\delta^{13}C$ (*53*). (**C**) $\delta^{18}O$ (*62*). (**D**) Continental fragmentation index (*10*). (**E**) Sedimentary material (*63*). (**F**) Estimated Paleozoic P_{CO_2} (*58*, *59*). (**G**) Rescaled species-diversity trajectory of the present study.

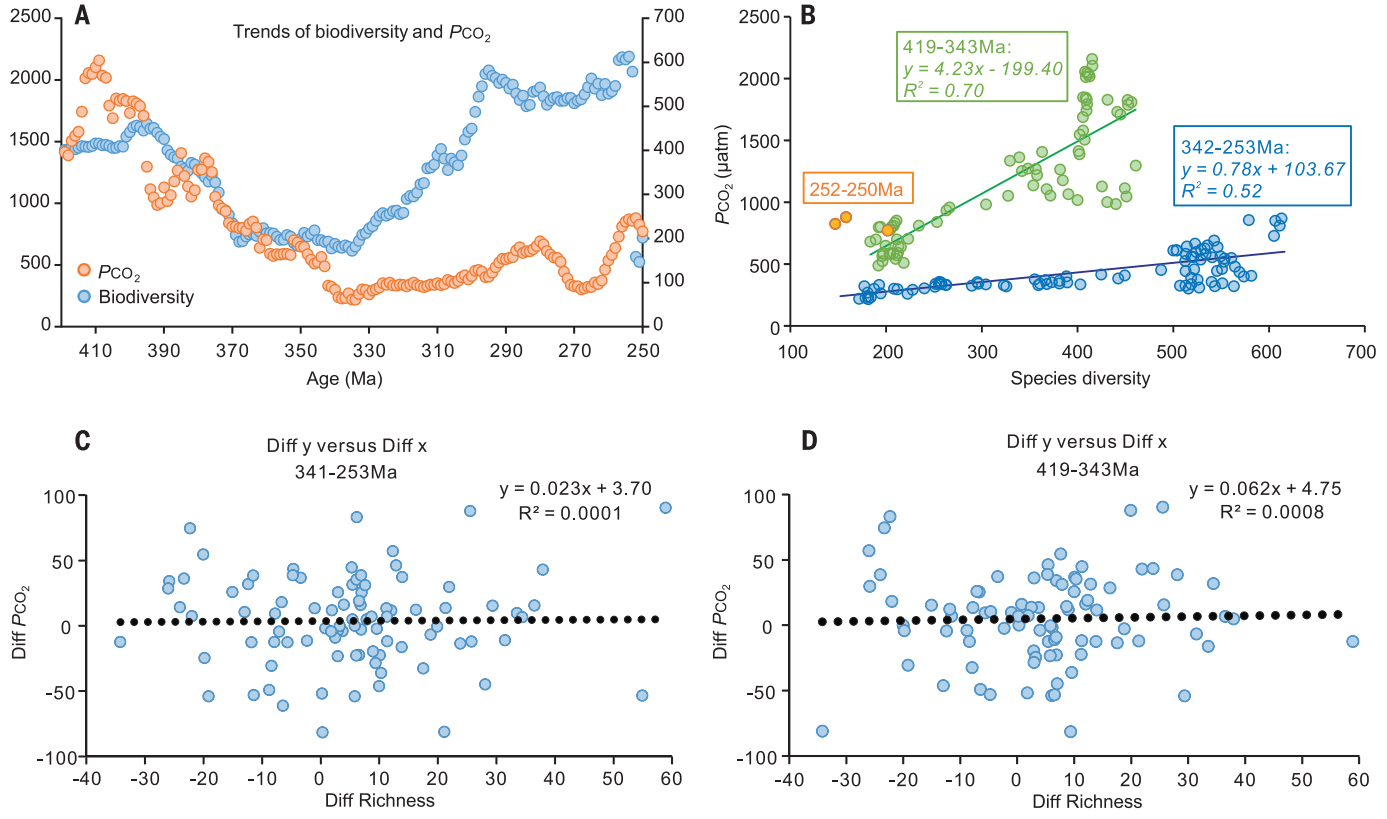


Fig. 5. Correlation analysis between species diversity and P_{CO_2} [data from (*58*, *59*)]. Results in (**A**) show strong secular trends of both diversity curves and Paleozoic P_{CO_2}. These secular trends generally decrease from 253 to 342 Ma and then increase from 343 to 419 Ma. The Pearson correlation analysis without considering the effects of the trends gives the results, coefficient of determination R^2 = 0.70, Student's *t*-value = 13.2, and *n* = 77 for the curve from 419 to 343 Ma and R^2 = 0.52, *t*-value = 9.72, and *n* = 90 for the other range between 342 and 253 Ma (**B**). To test for the effects of trends or autocorrelations in the time-series data, further correlation analysis was applied to the changes in species diversity and the changes in P_{CO_2} according to Yule's interpretation. The changes of values are calculated by first difference (Diff) method (**C** and **D**) in which each point is subtracted from the point that came before it. The results indicate that there is no significant correlation between the two curves of changes.

positively with increasing long-term diversity before the early Visean (*58*). After the post-Visean appearance of a more modern fauna, diversity increased more rapidly than the increase of P_{CO_2}. The similarity of these secular trends (Fig. 5, A and B) might imply that both P_{CO_2} and diversity were responding to a common set of driving factors. However, the results obtained from further correlation analysis (Fig. 5, C and D, and fig. S13) do not show any significant correlation or causal association between P_{CO_2} and diversity. Similarly, other environmental proxies, (e.g., $^{87}Sr/^{86}Sr$ ratio, $\delta^{13}C$, $\delta^{18}O$, continental fragmentation index) do not exhibit compelling trends with diversity changes (fig. S12). However, such comparisons are hampered by the absence of long-term high-resolution environmental proxy data (Fig. 4 and fig. S12).

Evidence for a decline in continental block fragmentation from the Cambrian to the Triassic (*10*) is negatively related to the general increase in diversity (Fig. 4E). Changes in the $^{87}Sr/^{86}Sr$ ratio are related to continental breakup, oceanic midridge spreading, and more intense continental weathering. Comparison between the $^{87}Sr/^{86}Sr$ ratio profile and the diversity pattern shows that each of the Paleozoic diversity crises (end-Ordovician, Middle to Late Devonian, and end-Permian) coincides with transitions from low $^{87}Sr/^{86}Sr$ ratios to increasing values, suggesting a reduction in seafloor spreading, an increase of the continental weathering, or both. However, the GOBE is associated with a steady decline in the $^{87}Sr/^{86}Sr$ ratio, whereas during the Carboniferous–Permian Biodiversification Event, the $^{87}Sr/^{86}Sr$ ratios were high (Fig. 4A), indicating increased continental weathering (*44*, *60*).

Conclusions

The combination of our new Chinese data compilation, a new parallel computing implementation of CONOP.SAGA stratigraphic correlation algorithms, and the parallel processing power of the Tianhe II supercomputer have allowed the construction of a high-resolution composite species-diversity history with an average resolving power of 26.0 ± 14.9 kyr. Our results indicate that the coarse and uneven temporal resolutions used by previous summaries artificially influenced paleobiodiversity estimations. This analysis confirms the existence of end-Ordovician and end-Permian mass extinctions, a long-term Middle to Late Devonian diversity decline, and a markedly subdued Frasnian–Famennian event. Three biodiversification events are also evident in our results: (i) from late Cambrian to Middle Ordovician, (ii) in early Silurian, and (iii) from late Carboniferous to Cisuralian. The proposed mid-Carboniferous and end-Guadalupian "mass extinctions" are only evident as minor species-diversity fluctuations. A recent long-term P_{CO_2} reconstruction (*59*) shows similar secular trends with the diversity data between the Silurian and Early Triassic. However, correlation between these two types of time-series curves shows strong autocorrelation and definite cause-and-effect links between specific environmental factors and the diversity changes require more study because of the lack of long-term, high-resolution, environmental

proxy data and/or discrepancies between the dates assigned to our biodiversity curve and the environmental proxy curves.

REFERENCES AND NOTES

1. J. J. Sepkoski Jr., *A Compendium of Fossil Marine Animal Families* (Milwaukee Public Museum, ed. 2, 1992)
2. J. J. Sepkoski Jr., *A Compendium of Fossil Marine Animal Genera* (Paleontological Research Institution, 2002).
3. J. Alroy *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6261–6266 (2001).
4. J. Alroy *et al*., *Science* **321**, 97–100 (2008).
5. R. K. Bambach, *Annu. Rev. Earth Planet. Sci.* **34**, 127–155 (2006).
6. C. M. O. Rasmussen, B. Kröger, M. L. Nielsen, J. Colmenar, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7207–7213 (2019).
7. J. Alroy, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11536–11542 (2008).
8. B. Hannisdal, S. E. Peters, *Science* **334**, 1121–1124 (2011).
9. S. E. Peters, M. Foote, *Paleobiology* **27**, 583–601 (2001).
10. A. Zaffos, S. Finnegan, S. E. Peters, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5653–5658 (2017).
11. M. J. Benton, Ed., *The Fossil Record 2* (Chapman & Hall, 1993).
12. J. Alroy, *Paleontol. Soc. Pap.* **16**, 55–80 (2010).
13. O. H. Walliser, Ed., *Global Events and Event Stratigraphy in the Phanerozoic* (Springer, 1996).
14. P. M. Sadler, R. A. Cooper, M. J. Melchin, *Geol. Soc. Am. Bull.* **121**, 887–906 (2009).
15. J. X. Fan *et al*., *Newsl. Stratigr.* **46**, 111–136 (2013).
16. Materials and methods are available as supplementary materials.
17. A. J. Boucot, X. Chen, C. R. Scotese, R. J. Morley, "Phanerozoic paleoclimate: An atlas of lithologic indicators of climate" in G. J. Nichols, B. Ricketts, Eds., *SEPM Concepts in Sedimentology and Paleontology* (SEPM, Tulsa, 2013), vol. 11, p. 478.
18. J. Y. Rong, J. X. Fan, A. I. Miller, G. X. Li, *Geol. J.* **42**, 431–454 (2007).
19. P. M. Sadler, R. A. Cooper, "Best-fit intervals and consensus sequences: comparison of the resolving power of traditional biostratigraphy and computer-assisted correlation" in *High Resolution Approaches in Paleontology*, P. J. Harries, Ed. (Kluwer-Academic, 2003), pp. 49-94.
20. R. A. Cooper, P. M. Sadler, A. Munnecke, J. S. Crampton, *Geol. Mag.* **151**, 349–364 (2014).
21. Y. Wang *et al*., *Paleobiology* **40**, 113–129 (2014).
22. P. M. Sadler, *Annu. Rev. Earth Planet. Sci.* **32**, 187–213 (2004).
23. P. M. Sadler, R. A. Cooper, M. J. Melchin, *Proc. Yorks. Geol. Soc.* **58**, 329–343 (2011).
24. K. M. Cohen, S. C. Finney, P. L. Gibbard, J. X. Fan, *Episodes* **36**, 199–204 (2013).
25. S. Z. Shen, J. Y. Rong, Eds., *New Advances in the Integrative Stratigraphy and Timescale of China*, vol. 62 of *Science China Earth Sciences* (Science China Press, 2019).
26. M. Foote, *Paleobiology* **26**, 578–605 (2000).
27. M. R. Saltzman *et al*., *J. Sediment. Res.* **74**, 366–377 (2004).
28. T. Servais, D. A. T. Harper, *Lethaia* **51**, 151–164 (2018).
29. J. Y. Rong, B. Huang, R. B. Zhan, D. A. T. Happer, Latest Ordovician and earliest Silurian brachiopods succeeding the Hirnantia fauna in south-east China. *Spec. Pap. Palaeontology* **90**, 5–142 (2013).
30. G. X. Wang, R. B. Zhan, I. G. Percival, *Earth Sci. Rev.* **192**, 15–33 (2019).
31. D. M. Raup, J. J. Sepkoski Jr., *Science* **215**, 1501–1503 (1982).
32. R. K. Bambach, A. H. Knoll, S. C. Wang, *Paleobiology* **30**, 522–542 (2004).
33. M. E. Clapham, S. Z. Shen, D. J. Bottjer, *Paleobiology* **35**, 32–50 (2009).
34. S. D. Burgess, S. Bowring, S. Z. Shen, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3316–3321 (2014).
35. J. J. Sepkoski Jr., *Paleobiology* **7**, 36–53 (1981).
36. S. Z. Shen, H. Zhang, W. Z. Li, L. Mu, J. F. Xie, *Geol. J.* **41**, 345–361 (2006).
37. A. L. Cárdenas, P. J. Harries, *Nat. Geosci.* **3**, 430–434 (2010).
38. D. H. Erwin, *Curr. Biol.* **19**, R575–R583 (2009).
39. P. J. Mayhew, M. A. Bell, T. G. Benton, A. J. McGowan, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15141–15145 (2012).
40. P. J. Mayhew, G. B. Jenkins, T. G. Benton, *Proc. Biol. Sci.* **275**, 47–53 (2008).
41. I. P. Montañez *et al*., *Science* **315**, 87–91 (2007).
42. J. A. Trotter, I. S. Williams, C. R. Barnes, C. Lécuyer, R. S. Nicoll, *Science* **321**, 550–554 (2008).
43. J. A. Trotter, I. S. Williams, C. R. Barnes, P. Männik, A. Simpson, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **443**, 34–48 (2016).
44. J. T. Chen, I. P. Montañez, Y. P. Qi, S. Z. Shen, X. D. Wang, *Geology* **46**, 395–398 (2018).
45. B. Chen *et al*., *Gondwana Res.* **24**, 77–89 (2013).
46. A. Delabroye, M. Vecoli, *Earth Sci. Rev.* **98**, 269–282 (2010).
47. S. Finnegan, N. A. Heim, S. E. Peters, W. W. Fischer, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6829–6834 (2012).
48. M. M. Joachimski *et al*., *Earth Planet. Sci. Lett.* **284**, 599–609 (2009).
49. C. Huang, M. M. Joachimski, Y. M. Gong, *Earth Planet. Sci. Lett.* **495**, 174–184 (2018).
50. J. Chen *et al*., *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **448**, 26–38 (2016).
51. M. M. Joachimski *et al*., *Geology* **40**, 195–198 (2012).
52. Y. Sun *et al*., *Science* **338**, 366–370 (2012).
53. M. R. Saltzman, E. Thomas, "Carbon isotope stratigraphy" in *The Geologic Time Scale 2012*, F. M. Gradstein, J. G. Ogg, M. D. Schmitz, G. M. Ogg, Eds. (Elsevier, 2012), vol. 1, pp. 207–232.
54. D. H. Rothman, *Sci. Adv.* **3**, e1700906 (2017).
55. T. G. Zhang, Y. N. Shen, R. B. Zhan, S. Z. Shen, X. Chen, *Geology* **37**, 299–302 (2009).
56. M. M. Joachimski, R. D. Pancost, K. H. Freeman, C. Ostertag-Henning, W. Buggisch, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **181**, 91–109 (2002).
57. R. A. Berner, Z. Kothavala, *Am. J. Sci.* **301**, 182–204 (2001).
58. G. L. Foster, D. L. Royer, D. J. Lunt, *Nat. Commun.* **8**, 14845 (2017).
59. C. R. Witkowski, J. W. H. Weijers, B. Blais, S. Schouten, J. S. Sinninghe Damsté, *Sci. Adv.* **4**, eaat4556 (2018).
60. J. M. McArthur, R. J. Howarth, G. A. Shields, "Strotium isotope stratigraphy" in *The Geological Time Scale 2012*, F. M. Gradstein, J. G. Ogg, M. D. Schmitz, G. M. Ogg, Eds. (Elsevier, 2012), vol. 2, pp. 127–144.
61. J. X. Fan *et al.*, Data and code for: A high-resolution summary of Cambrian to Early Triassic marine invertebrate biodiversity, Version 3, Dryad (2019); https://datadryad.org/stash/dataset/doi:10.5061/dryad.0gb5mkkx1.
62. J. Veizer, A. Prokoph, *Earth Sci. Rev.* **146**, 92–104 (2015).
63. W. W. Hay *et al*., *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **240**, 3–46 (2006).

ACKNOWLEDGMENTS

We thank G. Hunt, M. Foote, and an anonymous reviewer for their helpful reviews, which greatly improved the manuscript. We also thank Z. Shen for improving Fig. 5. **Funding:** This research was supported by the Strategic Priority Research Programs of the Chinese Academy of Sciences (grant nos. XDB26000000, XDB18000000, and XDA19050100), the National Natural Science Foundation of China (grant nos. 41725007 and 41420104003), and the National Key R&D Program of China (grant no. SQ2018YFE020456). The data from Tibet are supported by the Second Tibetan Plateau Scientific Expedition and Research (grant no. 2019QZKK0706). This paper is a contribution to the IUGS "Deep-Time Digital Earth" Big Science Program. **Author contributions:** J.F. and S.S. designed the research. S.S. and J.F. wrote the main text and supplementary materials with inputs from D.H.E., N.M., P.M.S., and Q.C. and data contributions from all authors. J.F., X.H., J.Y., and P.M.S. designed the CONOP.SAGA program for parallel computing. J.F. and S.S. performed most of the analyses. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data and code are available from the OneStratigraphy (http://onestratigraphy.ddeworld.org/download/525dcacb35e44dc0b273549324a26087) and Dryad (*61*) repositories.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/272/suppl/DC1
Materials and Methods
Figs. S1 to S13
Tables S1 to S3
References (*64*–*76*)

5 April 2019; accepted 22 November 2019
10.1126/science.aax4953

REPORT

SOLAR PHYSICS

Decay of the coronal magnetic field can release sufficient energy to power a solar flare

Gregory D. Fleishman[1]*, Dale E. Gary[1], Bin Chen[1], Natsuha Kuroda[2,3], Sijie Yu[1], Gelu M. Nita[1]

Solar flares are powered by a rapid release of energy in the solar corona, thought to be produced by the decay of the coronal magnetic field strength. Direct quantitative measurements of the evolving magnetic field strength are required to test this. We report microwave observations of a solar flare, showing spatial and temporal changes in the coronal magnetic field. The field decays at a rate of ~5 Gauss per second for 2 minutes, as measured within a flare subvolume of ~10^{28} cubic centimeters. This fast rate of decay implies a sufficiently strong electric field to account for the particle acceleration that produces the microwave emission. The decrease in stored magnetic energy is enough to power the solar flare, including the associated eruption, particle acceleration, and plasma heating.

The solar corona sometimes exhibits an explosive release of the energy stored in magnetized plasma, which drives phenomena such as solar flares (*1*–*5*). The standard model of solar flares (*6*–*9*) posits that they are powered by magnetic energy stored in the solar corona and released (dissipated into other forms) through magnetic reconnection (*10*)—a reconfiguration of the magnetic field topology toward a state of lower magnetic energy. Changes in the coronal magnetic field during a flare or other large-scale eruption have been quantified only indirectly, for example (*11*), from extrapolations of the magnetic field measured at the photosphere—the surface layer of the Sun seen in white light. Although this method can quantify the modest magnetic energy transfer of ~10%, it is known to suffer from many shortcomings (*12*). The extrapolation approach does not allow the dynamic local changes of the magnetic field to be quantified at time scales short enough to characterize the flare energy release.

We report observations (*13*) of a large solar flare—one of several that occurred in September 2017. The partially occulted eruptive flare occurred in active region (AR) 12673, at heliographic coordinates 9° south, 91° west (Fig. 1A), on 10 September 2017. This event exhibits the main ingredients of the standard flare model, including a cusplike structure of nested magnetic loops that evolves upward at a speed of ~30 km s^{-1} and an apparent current sheet (Fig. 1A) (*14*–*16*). This eruptive flare was widely observed at many wavelengths (*14*–*19*). Estimates of the kinetic, thermal, and nonthermal energies released in the flare are available from complementary approaches and datasets, whereas the dominant magnetic energy has only been estimated indirectly (*20*). Figure 1 shows context information for the flare, including the microwave images that we observed using the Expanded Owens Valley Solar Array (EOVSA) (*21*) in 26 microwave bands in the range of 3.4 to 15.9 GHz (*13*).

We produced magnetic field maps from these observations (*13*), examples of which are shown in Fig. 2. The full sequence is shown in movie S1. They show strong variation between maps, demonstrating the fast evolution of the coronal magnetic field strength B. The magnetic field strength decays quickly at the cusp region; away from that region, the field also decays but more slowly.

To quantify this decay, Fig. 3 shows the time evolution of the flaring coronal magnetic field at two locations marked in Fig. 2. Both locations exhibit a decay in the magnetic field strength but with different timing. One location shows a decay of the magnetic field from ~600 to ~200 G over ~1 min, a magnetic field decay rate of $|\dot{B}| \approx 6.6\ \mathrm{G\ s^{-1}}$. The decay ends at about 15:58 Coordinated Universal Time (UTC), after which the magnetic field at this location remains roughly constant. By contrast, the location within a larger and higher collapsing loop, marked in Fig. 1A, experiences a longer decay, until roughly 16:00 UTC. In this location, the magnetic field decays from ~900 to ~250 G over ~2 min, a rate of $|\dot{B}| \approx 5.4\ \mathrm{G\ s^{-1}}$. The energy release suggested by this magnetic

[1]Center for Solar Terrestrial Research, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA. [2]University Corporation for Atmospheric Research, Boulder, CO 80307, USA. [3]Space Science Division, Code 7684, Naval Research Laboratory, Washington, DC 20375, USA.
*Corresponding author. Email: gfleishm@njit.edu

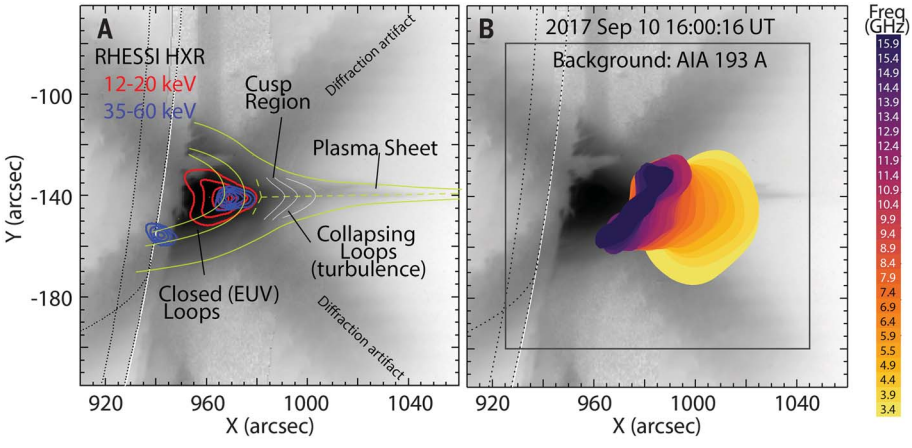


Fig. 1. Multiwavelength observations of the class X8 flare on 10 September 2017. (**A**) An EUV image (193 Å) with inverted brightness overlain with contours outlining the thermal (red contours) and nonthermal (blue contours) hard x-ray (HXR) emission (*16*). The green and white lines are a schematic drawing of the plasma sheet (the current sheet, according to the standard solar flare model), closed and collapsing (newly reconnected) loops, and the cusp region, where the fastest evolution of the magnetic field takes place. Only one of the loop foot points (the southern one) is located on the visible side of the disk, whereas the other is located behind the limb (occulted by the Sun). The thin white curve shows the solar surface (photosphere). The dotted black lines indicate the solar coordinate grid marked at 5° intervals. X and Y are the Cartesian coordinates with the coordinate center adopted in the center of the solar disk. RHESSI, Reuven Ramaty High Energy Solar Spectroscopic Imager. (**B**) The same image as (A), overlain with the microwave observations taken with EOVSA. The colored regions indicate the ≥50% brightness areas corresponding to 26 frequencies from 3.4 to 15.9 GHz. The relationships among different data sources suggest that the microwave emission comes from the cusp region, outlining the newly reconnected collapsing field lines. The gray box outlines the region of corresponding magnetic field maps in Fig. 2. AIA, Atmospheric Imaging Assembly; Freq., frequency.

field decay ends around the time of the peak microwave emission (16:00 UTC), which is consistent with the theoretical predictions of microwave emission arising from a population of trapped electrons (*16*).

We compare these measurements of the decaying coronal magnetic field with equilibrium models that are based on extrapolation of photospheric magnetic field measurements (*11*). The extrapolation requires corresponding photospheric vector magnetic field data, which are not available for this partially occulted event. However, magnetic field models are available (*22*) for this AR a few days earlier, on 6 September 2017, when this AR could be seen more face-on. These models found that the strongest magnetic field in the corona at the height of 30 Mm was ~200 G [figure 5 in (*22*)]. Our measurements at that height match this model value at the end of the time range analyzed, after the decay of the magnetic field is over, which implies that the magnetic field in the flare evolves toward an equilibrium state. However, the much stronger values observed earlier in the flare are several times as high as the equilibrium values. This indicates that a dynamic, transient magnetic field was lifted up from lower heights by the eruption process. This redistribution of the strong magnetic field—originally located low in the corona—over a much larger coronal volume during the flare might power the solar flare and associated eruption.

The Faraday equation is $\dot{\mathbf{B}} = -c\nabla \times \mathbf{E}$ (where $\mathbf{B}$ and $\mathbf{E}$ are the magnetic and electric field vectors, respectively, and c is the speed of light), which requires that an electric field be associated with the observed decay in magnetic field strength. Estimating $|\nabla \times \mathbf{E}|$ as E/R, where R is the scale of nonuniformity at the cusp region, and adopting representative values $\dot{B} \approx 5\,\mathrm{G\,s^{-1}}$ and $R \approx 3.65 \times 10^8$ cm—equivalent to 5 arc sec on the Sun, which is the size of the smallest coherent structures in the magnetic maps [as well as in extreme ultraviolet (EUV) images (*15*)]—we find $E \sim 20\,\mathrm{V\,cm^{-1}}$. For comparison, the Dreicer field, which demarcates regimes of the steady electric current and free runaway of the plasma electrons (*23*), is $E_D \sim 10^{-4}\,\mathrm{V\,cm^{-1}}$. This field strength E is consistent, within an order of magnitude, with available indirect estimates (*10*, *24*, *25*) for other large flares. Our choice of the scale $R \approx 5$ arc sec is ~10 to 20% of the cusp size and represents the macroscopic structure. It does not preclude the existence of smaller-scale structures and the proportionally smaller electric fields associated with them. However, the derived electric field remains above the Dreicer value for any scales $\gtrsim 5 \times 10^{-5}$ arc sec.

The decrease in magnetic energy at the cusp region must be associated with a conversion of that energy into other forms. An energy source is needed at the cusp region of this event to account for its enhanced temperature (*15*). The magnetic-field–aligned component of our inferred electric field should accelerate particles as required to power the microwave emission and could drive the observed enhanced heating at the cusp region (*15*).

The decay in magnetic field strength implies advection and/or diffusion of the magnetic field, which can be estimated using the induction equation, $\dot{\mathbf{B}} = \nabla \times [v \times \mathbf{B}] + \nu\nabla^2\mathbf{B}$, where v is the plasma velocity and ν is the magnetic diffusivity. The advection term, $\nabla \times [v \times \mathbf{B}]$, can easily account for the magnetic field variation during the phase of apparent upward motion of the arclike structures in the magnetic field maps from 15:57:00 to 15:59:25 UTC (movie S1). However, at later times these arclike structures fade without moving, which implies dissipation of the magnetic field, not advection. We estimate the magnetic diffusivity ν required to drive the apparent decay rate ($|\dot{B}| \approx 5\,\mathrm{G\,s^{-1}}$) of the magnetic field ($B \approx 600$ G) as $\nu \sim R^2\,\dot{B}/B \sim 10^{15}\,\mathrm{cm^2\,s^{-1}}$. This value of ν is much larger than the magnetic diffusivity due to Coulomb collisions (*26*). It can only be provided by turbulent magnetic diffusion, which appears when a large fraction of the velocity v entering the induction equation is fluctuating [turbulent-like (*27*)], rather than steady. Averaging the induction equation over the random velocity field results in a renormalization of the magnetic diffusivity coefficient (*26*) such that $\nu \sim u\,R/3$, where u is the typical turbulent velocity. Nonthermal turbulent velocities of $u \sim 100\,\mathrm{km\,s^{-1}}$ were measured at the cusp region in this flare (*15*), which implies $\nu \sim u\,R/3 \sim 1.2 \times 10^{15}\,\mathrm{cm^2\,s^{-1}}$, in agreement with the estimate obtained from $\dot{B}$ above.

Figure 3B shows the evolution of the mean magnetic energy density (*13*) in this region. Over ~1.5 min, the magnetic energy density decays at a rate of ~200 erg $\mathrm{cm^{-3}\,s^{-1}}$, losing ~80% of the magnetic energy available in this area. The net decrease of the magnetic energy in the adopted nominal flare volume of $10^{28}\,\mathrm{cm^3}$, which corresponds to a source with a linear scale of ~20 Mm (Fig. 1), is $\sim 2 \times 10^{32}$ erg.

We compare this reduction in magnetic energy density with the energy density of nonthermal electrons (microwave diagnostics provides the instantaneous electron and energy densities, unlike x-ray diagnostics, which provides electron and energy flux), which we compute from the same data (*13*). Figure 3B

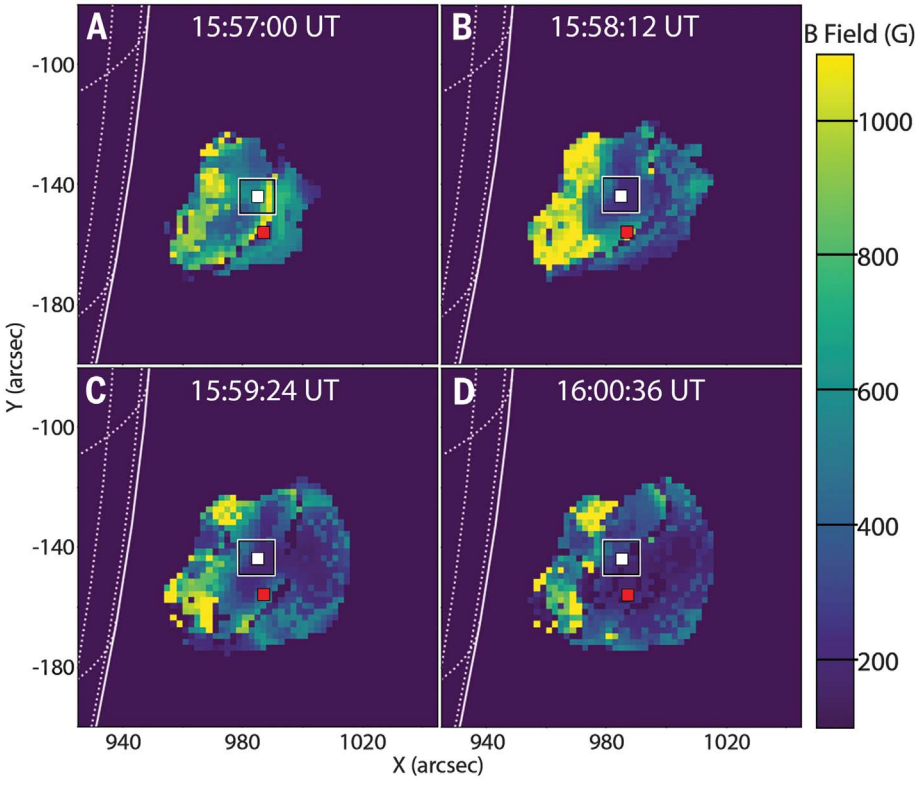


Fig. 2. Evolving maps of the coronal magnetic field. (**A** to **D**) Four coronal magnetic field maps derived for the 10 September 2017 flare, separated by 72 s. Apparent upward motion of the radio source and looplike structures in the magnetic field maps is visible in panels (A) to (C), showing the spreading of the reconnection process upward. Red and white squares, and the empty white box, correspond to locations shown in Fig. 3. The solar coordinate grid and the solar photosphere are shown by the white dotted and solid lines, respectively. Movie S1 shows an animated version of this figure.

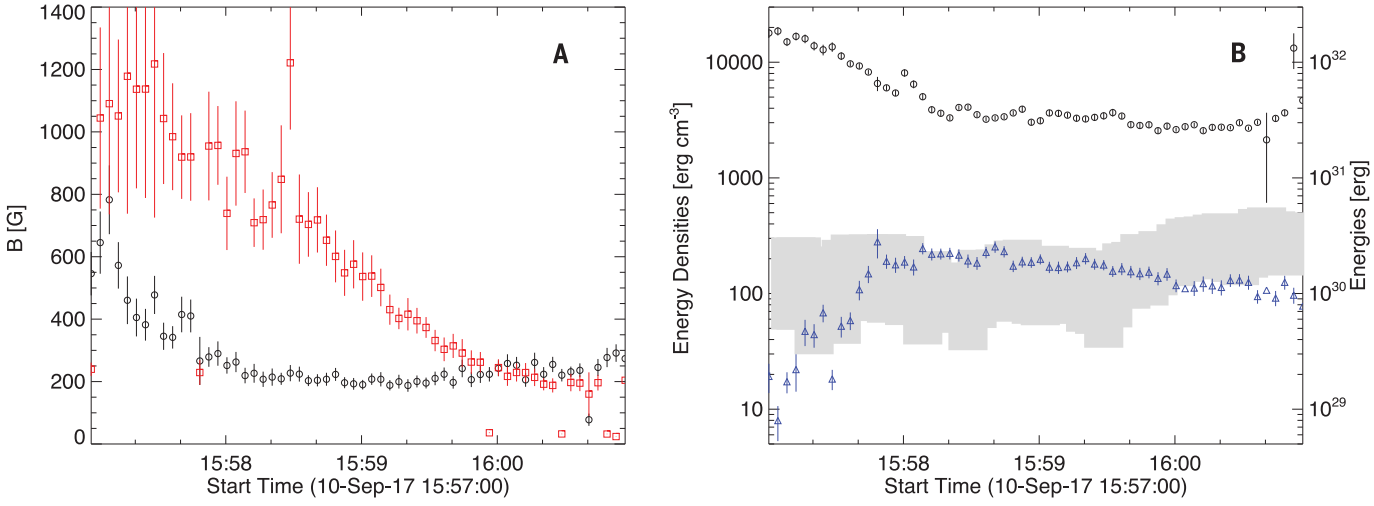


Fig. 3. Evolution of the magnetic field and magnetic energy. (**A**) Evolution of the magnetic field at two locations shown in Fig. 2. The red and black symbols show the data from the red and white squares, respectively. Black circles show decay of the magnetic field from 15:57 to 15:58 UTC, remaining roughly constant after that. The red squares, which correspond to a higher member of the system of nested loops, show a similar decay lasting 2 min longer, coinciding with the apparent upward motion of the EUV loops (Fig. 2). Error bars show 1σ uncertainties. (**B**) The mean magnetic energy density w_B (black circles) and the mean energy density of nonthermal electrons w_{nth} (blue triangles), both computed within the white box shown in Fig. 2. The shaded gray area indicates our estimated range of thermal energy density computed in (*13*). The right axis shows the corresponding total energy, assuming a flare volume of 10^{28} cm^3. Error bars show 1σ uncertainties.

shows the evolution of the (lower bound of the) nonthermal energy density, assuming a power-law distribution of electrons with a low-energy cutoff of E_{min} = 20 keV (*13*). Although this nonthermal energy constitutes only a few percent of the decrease in magnetic energy, it appears that the main decay of the magnetic energy is correlated in time with the increase of the nonthermal electron energy. This result implies that a direct energy conversion occurs in this region. The thermal energy density and kinetic energy density of turbulent motions cannot be estimated from the microwave diagnostics alone but require additional inputs based on available EUV diagnostics (*15*). The estimated thermal energy density (*13*) is also shown in Fig. 3B. It overlaps with the lower bound of the nonthermal energy density obtained above. The kinetic energy associated with random motions of the plasma is two orders of magnitude lower than the thermal energy. This implies that we observe a region at the cusp location—where available magnetic energy is converted to other forms of flare energy—which occurs below, but not within, the reconnection current sheet. The observed release of the magnetic energy is sufficient to power all other observed forms of energy in the flare.

Our observations quantify the coronal magnetic energy at the flare site and establish exactly where and how fast it is released. Our findings provide a quantitative observation of energy transformation in a solar flare, linking the thermal and nonthermal energy components to the associated magnetic energy release.

REFERENCES AND NOTES

1. T. Yokoyama, K. Shibata, *Nature* **375**, 42–44 (1995).
2. B. Kliem, T. Török, *Phys. Rev. Lett.* **96**, 255002 (2006).
3. K. Shibata *et al.*, *Science* **318**, 1591–1594 (2007).
4. S. Scaringi, T. J. Maccarone, C. D'Angelo, C. Knigge, P. J. Groot, *Nature* **552**, 210–213 (2017).
5. P. F. Wyper, S. K. Antiochos, C. R. DeVore, *Nature* **544**, 452–455 (2017).
6. H. Carmichael, in *The Physics of Solar Flares, Proceedings of the AAS-NASA Symposium held 28-30 October, 1963 at the Goddard Space Flight Center, Greenbelt, MD*, W. N. Hess, Ed. (National Aeronautics and Space Administration, 1964), vol. 50, p. 451.
7. P. A. Sturrock, *Nature* **211**, 695–697 (1966).
8. T. Hirayama, *Sol. Phys.* **34**, 323–338 (1974).
9. R. A. Kopp, G. W. Pneuman, *Sol. Phys.* **50**, 85 (1976).
10. E. Priest, T. Forbes, *Magnetic Reconnection* (Cambridge Univ. Press, 2007).
11. X. Sun *et al.*, *Astrophys. J.* **748**, 77 (2012).
12. M. L. DeRosa *et al.*, *Astrophys. J.* **696**, 1780–1791 (2009).
13. Materials and methods are available as supplementary materials.
14. Y. Li *et al.*, *Astrophys. J.* **853**, L15 (2018).
15. H. P. Warren *et al.*, *Astrophys. J.* **854**, 122 (2018).
16. D. E. Gary *et al.*, *Astrophys. J.* **863**, 83 (2018).
17. N. Omodei, M. Pesce-Rollins, F. Longo, A. Allafort, S. Krucker, *Astrophys. J.* **865**, L7 (2018).
18. D. M. Long *et al.*, *Astrophys. J.* **855**, 74 (2018).
19. G. A. Doschek *et al.*, *Astrophys. J.* **853**, 178 (2018).
20. D. Longcope, J. Unverferth, C. Klein, M. McCarthy, E. Priest, *Astrophys. J.* **868**, 148 (2018).
21. G. M. Nita, J. Hickish, D. MacMahon, D. E. Gary, *J. Astron. Instrum.* **05**, 1641009 (2016).
22. S. A. Anfinogentov, A. G. Stupishin, I. I. Mysh'yakov, G. D. Fleishman, *Astrophys. J.* **880**, L29 (2019).
23. H. Dreicer, *Phys. Rev.* **115**, 238–249 (1959).
24. J. Qiu, D. E. Gary, G. D. Fleishman, *Sol. Phys.* **255**, 107–118 (2009).
25. G. D. Fleishman *et al.*, *Astrophys. J.* **822**, 71 (2016).
26. G. D. Fleishman, I. N. Toptygin, *Cosmic Electrodynamics: Electrodynamics and Magnetic Hydrodynamics of Cosmic Plasmas* (Springer 2013).
27. J. T. Dahlin, J. F. Drake, M. Swisdak, *Phys. Plasmas* **24**, 092110 (2017).

ACKNOWLEDGMENTS

We thank the scientists and engineers who helped design and build EOVSA, especially G. Hurford, S. White, J. McTiernan, W. Grammer, and K. Nelin. **Funding:** This work was supported in part by NSF grants AGS-1817277, AST-1910354, and AGS-1654382 and NASA grants NNX17AB82G, 80NSSC18K0667, 80NSSC19K0068, and 80NSSC18K1128 to the New Jersey Institute of Technology. N.K. was partially supported by the NASA Living With a Star Jack Eddy Postdoctoral Fellowship Program, administered by UCAR's Cooperative Programs for the Advancement of Earth System Science (CPAESS). **Author contributions:** G.D.F. developed the methodology, performed the model fitting, analyzed the results, and wrote the draft manuscript; D.E.G. led the construction and commissioning of the EOVSA and developed the observational strategy and calibration for microwave spectroscopy; B.C. developed the microwave spectral imaging and self-calibration strategy; N.K. wrote software used in the analysis and visualization; D.E.G., B.C., and S.Y. prepared the microwave observation data; S.Y. implemented the microwave imaging pipeline under the guidance of D.E.G. and B.C.; and G.M.N. developed software used in modeling and testing the spectral fitting methodology and developed the GSFIT spectral fitting package. All authors discussed the interpretation of the data, contributed scientific results, and helped prepare the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Raw EOVSA observational data used for this study are available at www.ovsa.njit.edu/fits/IDB/20170910/IDB20170910155625/. Fully processed EOVSA spectral imaging data (in IDL save format) are available at http://ovsa.njit.edu/publications/fleishman_ea_science_2019/data/. The microwave data fitting software, GSFIT, is available in the community-contributed SolarSoftWare repository, under the packages category, at www.lmsal.com/solarsoft/ssw/packages/gsfit/.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/278/suppl/DC1
Materials and Methods
Figs. S1 and S2
References (*28*–*33*)
Movies S1 and S2

13 April 2019; accepted 4 December 2019
10.1126/science.aax6874

ORGANIC CHEMISTRY

Nitromethane as a nitrogen donor in Schmidt-type formation of amides and nitriles

Jianzhong Liu[1], Cheng Zhang[1], Ziyao Zhang[1], Xiaojin Wen[1], Xiaodong Dou[1], Jialiang Wei[1], Xu Qiu[1], Song Song[1], Ning Jiao[1,2]*

The Schmidt reaction has been an efficient and widely used synthetic approach to amides and nitriles since its discovery in 1923. However, its application often entails the use of volatile, potentially explosive, and highly toxic azide reagents. Here, we report a sequence whereby triflic anhydride and formic and acetic acids activate the bulk chemical nitromethane to serve as a nitrogen donor in place of azides in Schmidt-like reactions. This protocol further expands the substrate scope to alkynes and simple alkyl benzenes for the preparation of amides and nitriles.

Amides and nitriles are broadly important compounds for the synthesis of materials, agrochemicals, and pharmaceuticals (*1*, *2*). The Schmidt reaction is one of the most efficient nitrogenation approaches to access amides and nitriles from aldehydes and ketones with HN_3 or alkyl azides (*3*–*12*). Discovered in 1923, it has been used extensively by synthetic chemists for almost a century. However, its reliance on volatile, highly toxic, and potentially explosive azide reagents leaves room for further development (*13*–*15*) (Fig. 1A). Despite the long-standing search for hydrazoic acid replacements, the use of azide remains prevalent (*16*–*18*). The Beckmann rearrangement starting from oxime substrates is a well-known alternative approach to secondary amides (*19*, *20*). However, mild Beckmann rearrangements are still rare (*21*). Some specialized active oxime substrates, strong protic acid promotion, or preactivation of the hydroxyl (e.g., by tosylation) were usually required in classic catalytic Beckmann rearrangements (*22*–*25*). Moreover, traditionally, aldoximes are rarely transformed into the corresponding nitriles and N-unsubstituted amides (*25*).

Ideally, if one bulk and common chemical could be activated and endowed with new reactivity (*26*–*35*), it would promote synthetic innovation and industrial development. Because of the strongly electron-withdrawing properties of the nitro group, nitromethane has often been used as a carbon pronucleophile in transformations such as the traditional Henry or nitroaldol reactions (*36*, *37*), cross-dehydrogenative coupling reactions (*38*), and

[1]State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China. [2]State Key Laboratory of Organometallic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China.
*Corresponding author. Email: jiaoning@pku.edu.cn

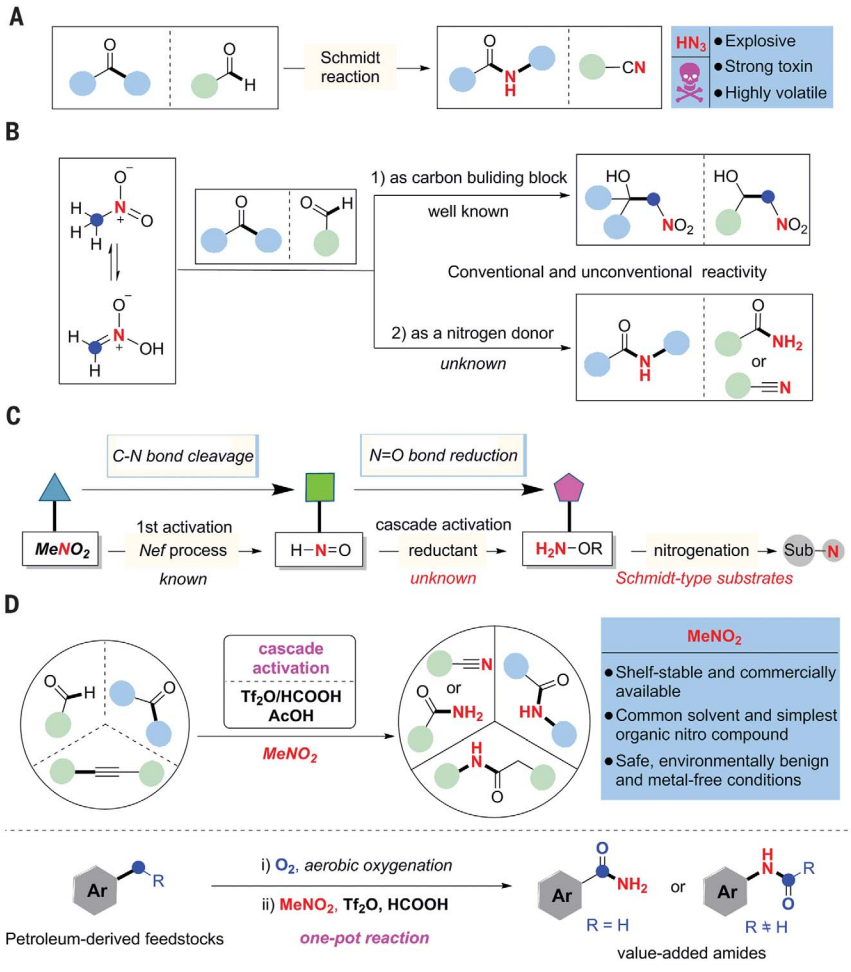


Fig. 1. Nitromethane activation for use in the Schmidt reaction. (**A**) Long-standing drawback of traditional Schmidt reaction with azide as the limiting reagent. (**B**) Reactivity patterns of nitromethane. (**C**) Proposed cascade activation strategy for the discovery of distinct reactivity of nitromethane inspired by the Nef process: activation of nitromethane to oxidation-state–matchable species endowed with Schmidt-type reactivity. (**D**) This work: direct nitrogenation of aldehydes, ketones, and alkynes, as well as the aerobic oxidative nitrogenation of alkylarenes with nitromethane.

Michael addition reactions (*39*) (Fig. 1B). By contrast, the reactivity of nitromethane as a nitrogen donor has been scarcely developed (*40*). It was reported that the electrophilic nitroxyl species (HNO) could be generated through the Nef process (*41*, *42*) (Fig. 1C); however, it is rather unstable because of its quick dimerization (rate constant $k = 8 \times 10^6$ $M^{-1}\,s^{-1}$) to produce inert N_2O and its oxidation by O_2 ($k = 3 \times 10^3$ $M^{-1}\,s^{-1}$) (*43*), which therefore limit its broad application. We hypothesized that a cascade reductive activation sequence could reduce the nitroxyl species (HNO) to a hydroxylamine derivative (NH_2OR), which would be more reductive than HNO, to

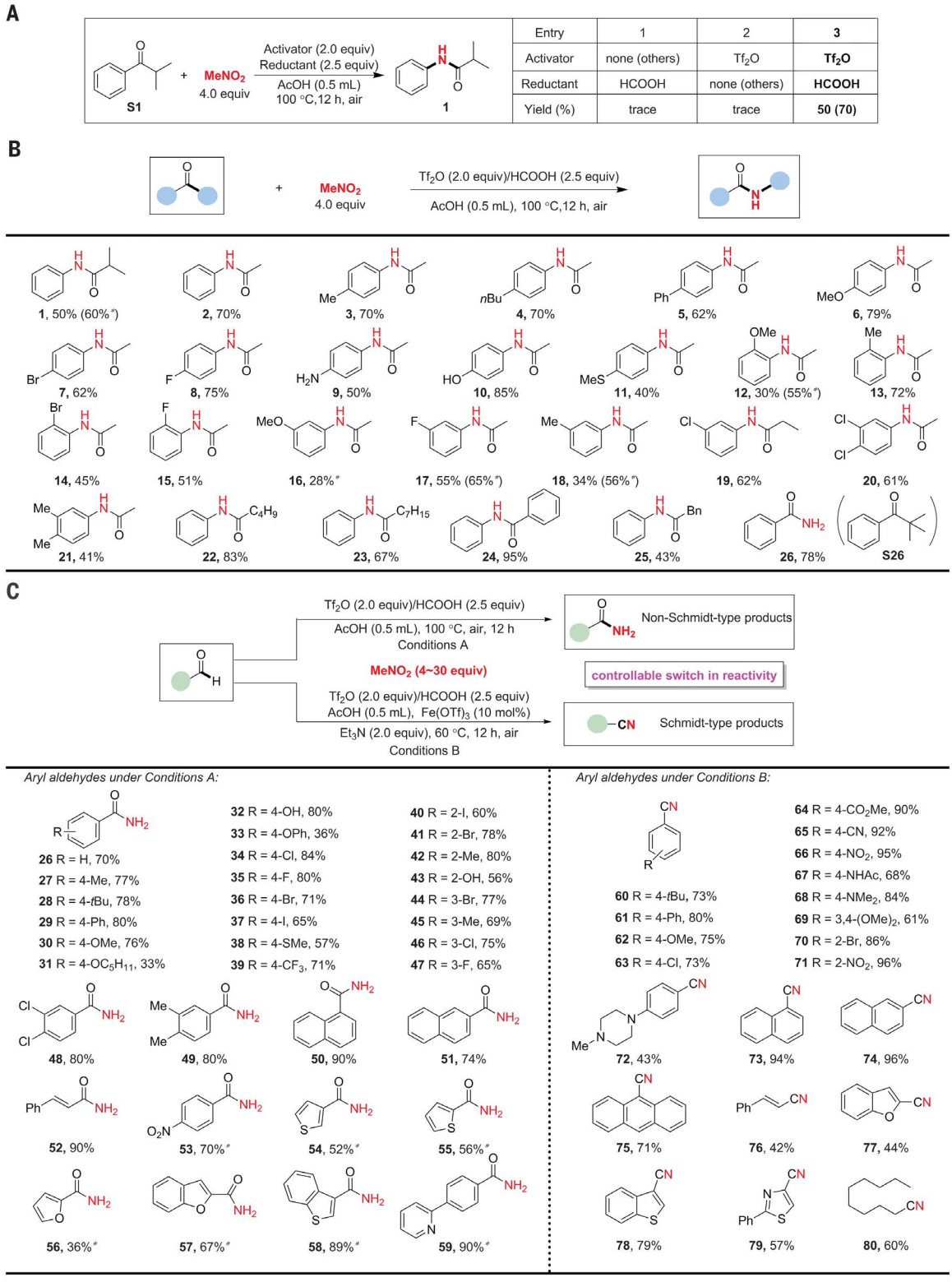


Fig. 2. Substrate scope of ketone and aldehyde derivatives. (**A**) Screening of the activator and reductant for the nitromethane activation. Reaction conditions: **S1** (0.3 mmol), $MeNO_2$ (1.2 mmol), activator (0.6 mmol), reductant (0.75 mmol) in AcOH (0.5 ml) at 100°C for 12 hours under air. The yield of amides in parentheses was determined by ^1H NMR analysis (details in supplementary materials). equiv, equivalent. (**B**) Substrate scope of ketones. (**C**) Substrate scope of aldehydes (details in supplementary materials). *$MeNO_2$ (9.0 mmol) was used. tBu, tertiary butyl; Me, methyl; Ph, phenyl.

facilitate the condensation and the subsequent rearrangement processes for Schmidt-type reactions (Fig. 1C). Thus, this cascade activation strategy would render simple bulk nitromethane a nitrogen donor akin to azides, but with considerably less associated hazard. We report here that nitromethane can indeed donate its nitrogen to aldehydes, ketones, alkynes, and even simple alkylbenzenes for the preparation of valuable amides or nitriles (Fig. 1D).

As a proof-of-principle study, we began our investigation by evaluating the Schmidt reaction with isobutyrophenone (**S1**) as substrate. After extensive screening (tables S1 and S2), we realized the targeted reactivity by tandem activation of the nitromethane with triflic anhydride (Tf_2O) and formic acid (HCOOH), and the desired amide product (**1**) was observed in 70% yield on the basis of integration of nuclear magnetic resonance (NMR) spectra with HCHO and TfOH as the by-products (Fig. 2A). Four equivalents of nitromethane were sufficient to complete this transformation, whereas lower loading of nitromethane decreased the efficiency. Acetic acid (AcOH) was also crucial for the high efficiency of this transformation (table S3).

As shown in Fig. 2B, numerous acetophenones were well tolerated in this process and provided the corresponding Schmidt-type amide products in moderate to good yields. In particular, a free amine group (**9**) and unprotected hydroxyl (**10**) proved compatible. The trialkyl-substituted acetophenone **S26** diverged from the reactivity pattern observed with the other substrates, providing the dealkylated free benzamide **26** in 78% yield. The reactions of unsymmetrical diaryl ketones demonstrate that electron-rich aromatic rings are more prone to migrating than electron-deficient aromatic rings (fig. S4). Notably, when we explored aldehyde substrates (Fig. 2C), the primary benzamides were obtained under the standard conditions with the formation of trace amount of benzonitriles (table S6). Considering that a Lewis acid may promote the rearrangement of the oxime intermediate to produce benzonitrile products as in the traditional Schmidt reaction, we explored the transformations in the presence of different Lewis acids (table S6). The reaction with the synergistic assistance of an iron(III) catalyst and triethylamine (Et_3N) base produced the corresponding nitriles in moderate to excellent yields (up to 96%) (Fig. 2C). We speculate that $Fe(OTf)_3$ and Et_3N serve as a Lewis acid and a base, respectively, to facilitate the elimination process for the formation of nitriles (table S6 and fig. S6). The aliphatic aldehyde (**80**) and (hetero)aryl aldehydes (**60** to **79**) were well tolerated. In the absence of the Lewis acid catalyst, the reaction afforded benzamide products (Fig. 2C). The competing

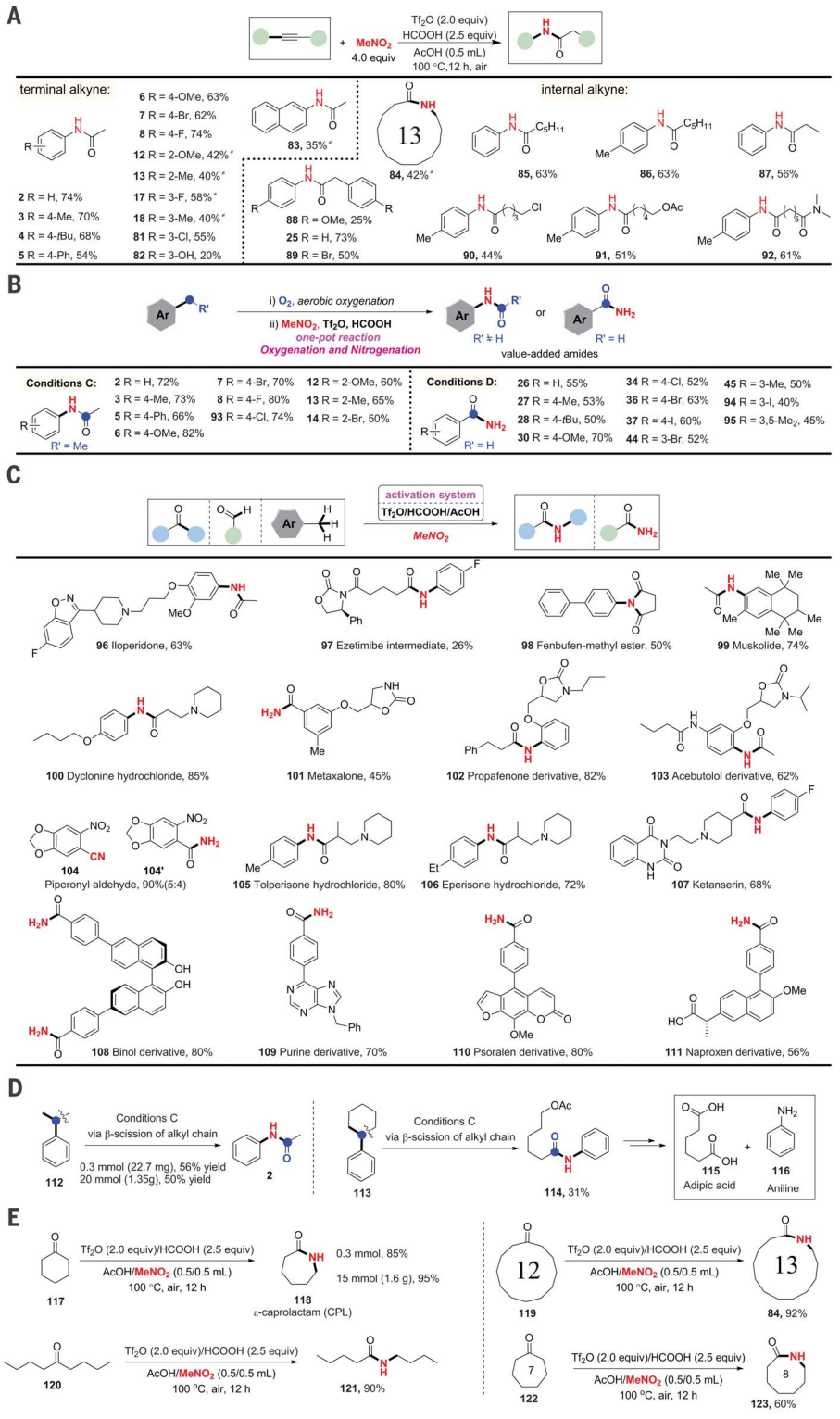


Fig. 3. Further synthetic applications. (**A**) Substrate scope of alkynes. (**B**) Nitrogenation of simple alkylarenes. Conditions C: ethylbenzene (0.3 mmol), $Co(OAc)_2{\cdot}4H_2O$ (0.015 mmol), *N*-hydroxyphthalimide (0.03 mmol) in AcOH (0.5 ml) were stirred at 80°C for 12 hours under 1 atm O_2; then, $MeNO_2$ (9.0 mmol), Tf_2O (0.6 mmol), HCOOH (0.75 mmol) were added and stirred at 100°C for 12 hours under air. Conditions D: methylbenzene (0.3 mmol), $Co(OAc)_2{\cdot}4H_2O$ (0.006 mmol), *N*-hydroxyphthalimide (0.03 mmol) in hexafluoroisopropanol (0.6 ml) were stirred at room temperature for 12 hours under 1 atm O_2; then, $MeNO_2$ (9.0 mmol), Tf_2O (0.6 mmol), HCOOH (0.75 mmol), AcOH (0.5 ml) were added and stirred at 100°C for 12 hours under air. (**C**) Late-stage modification of drugs or bioactive molecules. (**D**) Gram-scale reaction with cumene and cyclohexylbenzene. (**E**) CPL and macrocyclic lactam synthesis from cyclic ketones.

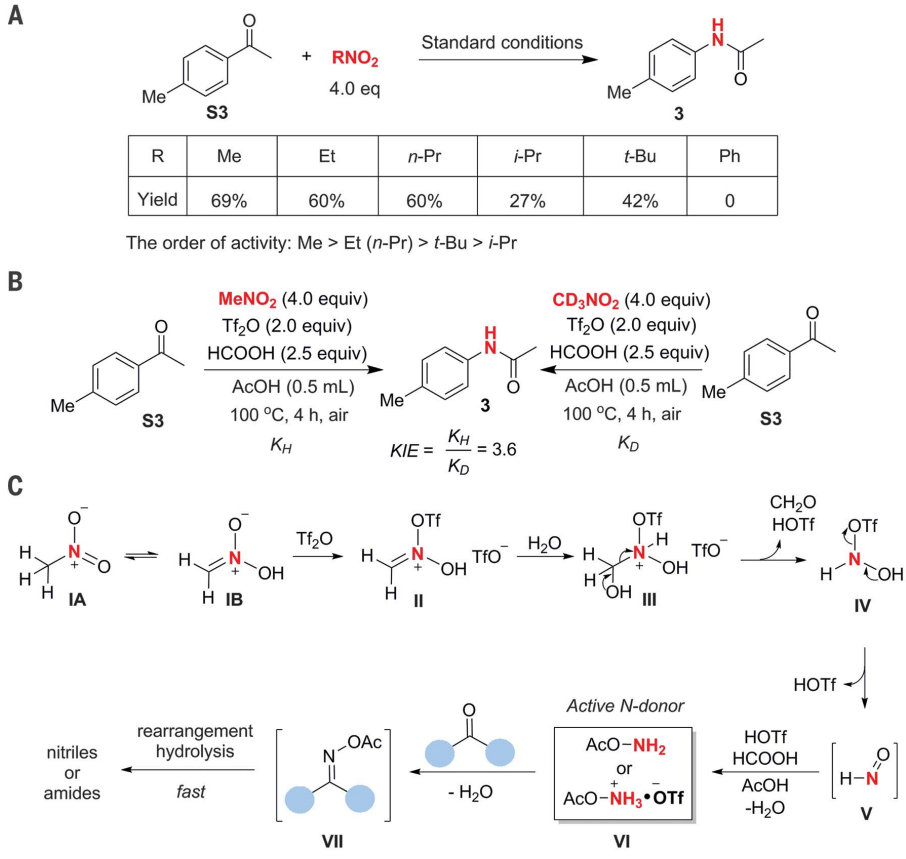


Fig. 4. Mechanistic experiments and proposed mechanism. (**A**) Structure-reactivity relationship of nitroalkanes. (**B**) KIE study. (**C**) Proposed mechanism. n-Pr, n-propyl; i-Pr, isopropyl.

aromatic nitro group in (**53**, **66**, and **71**) was not disturbed, highlighting the high chemoselectivity of the nitromethane activation protocol.

To further extend this strategy, we investigated simple alkyne substrate, which could produce ketones in situ by hydration (*44*) (Fig. 3A). A broad array of terminal alkynes could be efficiently converted to amide products. Moreover, unsymmetrical internal alkynes afforded the corresponding *N*-aryl amide products with high selectivity (**85** to **87** and **90** to **92**). The alkyl internal alkyne **84** also performed well. To clarify the nitrogenation chemistry of the alkynes, we conducted control experiments and in situ infrared experiments. These results indicate that the corresponding ketones are possible intermediates through alkyne hydration (fig. S5).

As part of an overarching goal of the nitrogenation of simple hydrocarbons (*45*–*48*), we next explored the tandem oxidative transformation of bulk commodity compounds such as alkylarenes (Fig. 3B). In this case, we tried to use the Co/NHPI/O_2 oxidative system (*49*) to produce the corresponding ketones and aldehydes in situ, which might enable the ensuing insertion of nitrogen by the present protocol. A variety of ethylbenzene and methylbenzene derivatives could produce the desired value-added amides by one-pot reaction with O_2 as environmentally benign oxidant and nitromethane as the nitrogen source for the nitrogen incorporation reaction. Neither purification nor isolation of the oxidized intermediate compounds was required, showing the robustness and the potential for industrial applications of this protocol.

Considering that ketone motifs are readily encountered in pharmaceutical compounds, the late-stage modification of such drugs and bioactive molecules would showcase the prospective utility of this protocol. The corresponding valuable amidated derivatives were delivered in moderate to excellent yields from structurally complex substrates containing carbonyl groups, including marketed drugs and their derivatives (**96** to **98**, **100** to **107**, and **110** and **111**) (Fig. 3C). Muskolide (**99**), piperonyl aldehyde (**104**), purine derivatives (**109**), and binol ligands (**108**) were also efficiently modified. The methyl group in metaxalone (**101**) could also be transformed into an amide group through the tandem oxygenation and nitrogenation sequence. The variety of tolerated functional groups further demonstrates the mildness and practicality of this protocol.

When cumene, a feedstock material of the Hock reaction for industrial phenol production, was subjected to the current protocol, amide **2** was obtained through oxidative β scission (*50*) in 50% yield at gram scale (Fig. 3D). Moreover, a similar reaction of cyclohexylbenzene produced amide **114**, which could be further transformed to the nylon 6.6 precursor adipic acid (**115**), as well as aniline (**116**) (Fig. 3D). ε-Caprolactam (CPL), the monomer of nylon 6, could also be efficiently synthesized from cyclohexanone by using this developed method. The products were isolated by direct recrystallization after a routine work-up procedure without column chromatography (Fig. 3E). The synthetically challenging macrocyclic lactam **84** was also easily prepared by this protocol. Moreover, long-chain aliphatic ketone **120** and 7-membered cycloketone **122** were also compatible with this transformation.

The alkyl substituent of the nitro group is necessary to this transformation, as shown in Fig. 4A: When the nitromethane was replaced with nitrobenzene, the reaction did not proceed. The reactivity dependence on alkyl groups was ordered as follows: primary carbon > tertiary carbon > secondary carbon (fig. S2). A measured kinetic isotope effect (KIE) of $k_H/k_D = 3.6$ suggests that the C(sp^3)–H bond cleavage of nitromethane might be involved in the rate-determining step of the reaction (Fig. 4B and fig. S3). Further study by ^{1}H NMR spectroscopy indicated that HCOOH is required for the second activation step (fig. S7). To identify the actual active N-donor species in this protocol, we carried out a high-resolution mass spectrometry (HRMS) capture experiment. Notably, an acetylated hydroxylamine (NH_2OAc) intermediate and its salt (NH_2OAc·HOTf) were detected by HRMS (fig. S8). Further control experiments suggest that these serve as the actual nitrogen donor of this system (supplementary materials). Meanwhile, these results demonstrate that the AcOH is also crucial in this transformation, playing a role not only as a solvent but also as a reagent. Inspired by the traditional Nef reaction and our findings, we propose the mechanism in Fig. 4C. Initially, nitromethane is activated by Tf_2O, forming the highly electrophilic imine species **II**. Subsequently, the reaction of intermediate **II** with H_2O forms hemiacetal intermediate **III**, which then undergoes intramolecular C–N bond cleavage to liberate HCHO (*41*, *42*) and **IV**. The active HNO species **V** forms after the elimination of HOTf from **IV**. Finally, the high oxidation state of HNO species **V** is selectively reduced by HCOOH to give an active N-donor species **VI** (NH_2OAc·HOTf or NH_2OAc), which was detected by HRMS. Overreduction is suppressed by the use of HCOOH as a reductant with moderate reducing capacity (*51*). The starting materials then react with active **VI** to form

the final products probably through the fast Beckmann-type rearrangement of the corresponding oximes **VII** (Fig. 4C).

REFERENCES AND NOTES

1. V. R. Pattabiraman, J. W. Bode, *Nature* **480**, 471–479 (2011).
2. J. Boström, D. G. Brown, R. J. Young, G. M. Keserü, *Nat. Rev. Drug Discov.* **17**, 709–727 (2018).
3. K. F. Schmidt, *Angew. Chem.* **36**, 511 (1923).
4. K. F. Schmidt, *Ber. Dtsch. Chem. Ges.* **57B**, 704–706 (1924).
5. H. Wolff, *Org. React.* **3**, 307–336 (1946).
6. P. A. S. Smith, *J. Am. Chem. Soc.* **70**, 320–323 (1948).
7. J. Aubé, G. L. Milligan, *J. Am. Chem. Soc.* **113**, 8965–8966 (1991).
8. R. V. Hoffman, J. M. Salvador, *Tetrahedron Lett.* **30**, 4207–4210 (1989).
9. R. V. Hoffman, J. M. Salvador, *J. Org. Chem.* **57**, 4487–4490 (1992).
10. W. H. Pearson, R. Walavalkar, J. M. Schkeryantz, W. Fang, J. D. Blickensdorf, *J. Am. Chem. Soc.* **115**, 10183–10194 (1993).
11. A. Wrobleski, T. C. Coombs, C. W. Huh, S.-W. Li, J. Aubé, *Org. React.* **78**, 307–336 (2012).
12. M. Szostak, J. Aubé, *Chem. Rev.* **113**, 5701–5765 (2013).
13. S. Bräse, C. Gil, K. Knepper, V. Zimmermann, *Angew. Chem. Int. Ed.* **44**, 5188–5240 (2005).
14. A. Hassner, M. Stern, H. E. Gottlieb, F. Frolow, *J. Org. Chem.* **55**, 2304–2306 (1990).
15. L. Marinescu, J. Thinggaard, I. B. Thomsen, M. Bols, *J. Org. Chem.* **68**, 9453–9455 (2003).
16. L. He, M. Wanunu, H.-S. Byun, R. Bittman, *J. Org. Chem.* **64**, 6049–6055 (1999).
17. C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **67**, 3057–3064 (2002).
18. K. C. K. Swamy, N. N. B. Kumar, E. Balaraman, K. V. P. P. Kumar, *Chem. Rev.* **109**, 2551–2651 (2009).
19. R. E. Gawley, *Org. React.* **35**, 1–61 (1988).
20. A. H. Blatt, *Chem. Rev.* **12**, 215–260 (1933).
21. X. Mo, T. D. R. Morgan, H. T. Ang, D. G. Hall, *J. Am. Chem. Soc.* **140**, 5264–5271 (2018).
22. H. Fujioka *et al*., *Chem. Pharm. Bull. (Tokyo)* **64**, 718–722 (2016).
23. A. Alhifthi, B. L. Harris, L. Goerigk, J. M. White, S. J. Williams, *Org. Biomol. Chem.* **15**, 10105–10115 (2017).
24. S. J. Touchette, E. M. Dunkley, L. L. Lowder, J. Wu, *Chem. Sci.* **10**, 7812–7815 (2019).
25. P. Crochet, V. Cadierno, *Chem. Commun.* **51**, 2495–2505 (2015).
26. A. J. A. Watson, J. M. J. Williams, *Science* **329**, 635–636 (2010).
27. J. R. Ludwig, P. M. Zimmerman, J. B. Gianino, C. S. Schindler, *Nature* **533**, 374–379 (2016).
28. D. Willcox *et al*., *Science* **354**, 851–857 (2016).
29. F. Juliá-Hernández, T. Moragas, J. Cornella, R. Martin, *Nature* **545**, 84–88 (2017).
30. L. Wang, J. M. Lear, S. M. Rafferty, S. C. Fosu, D. A. Nagib, *Science* **362**, 225–229 (2018).
31. H. Wang, X.-J. Dai, C.-J. Li, *Nat. Chem.* **9**, 374–378 (2017).
32. J. Jin, D. W. C. MacMillan, *Nature* **525**, 87–90 (2015).
33. Z. C. Litman, Y. Wang, H. Zhao, J. F. Hartwig, *Nature* **560**, 355–359 (2018).
34. M. Liu *et al*., *Chem* **5**, 858–867 (2019).
35. A. Hu, J.-J. Guo, H. Pan, Z. Zuo, *Science* **361**, 668–672 (2018).
36. G. Rosini, in *Comprehensive Organic Synthesis*, B. M. Trost, I. Fleming, Eds. (Pergamon, Oxford, 1991), pp. 321–340.
37. E. Jacobsen, in *The Nitro Group in Organic Synthesis*, N. Ono, Ed. (Wiley, New York, 2001), pp. 30–69.
38. Z. Li, D. S. Bohle, C.-J. Li, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8928–8933 (2006).
39. A. Noble, J. C. Anderson, *Chem. Rev.* **113**, 2887–2939 (2013).
40. Y. Yan *et al*., *Adv. Synth. Catal.* **358**, 212–217 (2016).
41. W. E. Noland, *Chem. Rev.* **55**, 137–155 (1955).
42. J. E. McMurry, J. Melton, *J. Org. Chem.* **38**, 4367–4373 (1973).
43. F. Doctorovich, D. E. Bikiel, J. Pellegrino, S. A. Suarez, M. A. Martí, in *Progress in Inorganic Chemistry*, K. D. Karlin, Ed. (Wiley, 2014), vol. 58, pp. 145–183.
44. L. Chen, C.-J. Li, *Adv. Synth. Catal.* **348**, 1459–1484 (2006).
45. M. P. Paudyal *et al*., *Science* **353**, 1144–1147 (2016).
46. L. Legnani, G. Prina-Cerai, T. Delcaillau, S. Willems, B. Morandi, *Science* **362**, 434–439 (2018).
47. J. Liu *et al*., *Nat. Chem.* **11**, 71–77 (2019).
48. T. Wang, N. Jiao, *Acc. Chem. Res.* **47**, 1137–1145 (2014).
49. F. Recupero, C. Punta, *Chem. Rev.* **107**, 3800–3842 (2007).
50. T. S. S. Rao, S. Awasthi, *E-J. Chem.* **4**, 1–13 (2007).
51. L. Knorr, H. Lange, *Ber. Dtsch. Chem. Ges.* **35**, 2998–3008 (1902).

ACKNOWLEDGMENTS

We thank X. Shi and X. Yang at State Key Laboratory of Natural and Biomimetic Drugs for help with the HRMS experiments and the in situ infrared experiments. **Funding:** We acknowledge the NSFC (nos. 21632001, 21772002, 81821004), the National Basic Research Program of China (973 Program) (no. 2015CB856600), and the Drug Innovation Major Project (2018ZX09711-001) for financial support. **Author contributions:** J.L. and N.J. conceived and designed the experiments. J.L. and C.Z. carried out most experiments; J.L., C.Z., Z.Z., X.W., X.D., J.W., X.Q., S.S., and N.J. analyzed data. J.L. and N.J. wrote the paper. N.J. directed the project. **Competing interests:** N.J., J.L., and C.Z. are inventors on patent application CN201910943969.5, submitted by Peking University, that covers the activation system and its application in nitrogenation reactions. All authors declare no other competing financial interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/281/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S8
Tables S1 to S6
Spectral Data
References (*52*–*109*)

31 July 2019; accepted 20 November 2019
Published online 5 December 2019
10.1126/science.aay9501

QUANTUM CRITICALITY

Singular charge fluctuations at a magnetic quantum critical point

L. Prochaska[1]*, X. Li[2]*†, D. C. MacFarland[1,3]*‡, A. M. Andrews[3], M. Bonta[4], E. F. Bianco[5]§, S. Yazdi[6]¶, W. Schrenk[7], H. Detz[7]#, A. Limbeck[4], Q. Si[8], E. Ringe[6]**, G. Strasser[3,7], J. Kono[2,6,8], S. Paschen[1,8]††

Strange metal behavior is ubiquitous in correlated materials, ranging from cuprate superconductors to bilayer graphene, and may arise from physics beyond the quantum fluctuations of a Landau order parameter. In quantum-critical heavy-fermion antiferromagnets, such physics may be realized as critical Kondo entanglement of spin and charge and probed with optical conductivity. We present terahertz time-domain transmission spectroscopy on molecular beam epitaxy–grown thin films of $YbRh_2Si_2$, a model strange-metal compound. We observed frequency over temperature scaling of the optical conductivity as a hallmark of beyond-Landau quantum criticality. Our discovery suggests that critical charge fluctuations play a central role in the strange metal behavior, elucidating one of the long-standing mysteries of correlated quantum matter.

Quantum critical behavior as prescribed by the Landau framework of order parameter fluctuations (*1*, *2*) has been clearly identified in insulating quantum magnets such as $LiHoF_4$ (*3*) and $TlCuCl_3$ (*4*). In strongly correlated metals, however, this framework often fails. In the strange-metal (*5*) regime of various correlated systems (*6*), electronic localization-delocalization transitions have been reported (*7*–*14*), and it is an outstanding question whether they are a key ingredient of beyond-Landau quantum criticality. To make progress, it is essential to study the dynamics of charge carriers in a suitable setting.

We chose the heavy fermion metal $YbRh_2Si_2$ (*15*) for our investigation because it has a well-defined quantum critical point (*15*, *16*) and shows evidence for an electron localization-delocalization transition (*7*, *8*) in its strange-metal regime. An ideal tool to study such properties is optical conductivity measurements in the relevant frequency window, which is typically the terahertz range and below for heavy fermion systems. However, such measurements are challenging on bulk samples because the Kramers-Kronig transformation to extract the real and imaginary parts of the optical conductivity from reflectivity measurements introduces substantial uncertainty at low frequencies (*17*). Thus, we resorted to a different approach: We performed terahertz time-domain transmission spectroscopy experiments on thin films of $YbRh_2Si_2$ grown by means of molecular beam epitaxy (MBE). Our measurements reveal ω/T scaling of the optical conductivity, where ω is the (angular) frequency and T is the temperature, elucidating the mechanism for strange-metal phenomena.

To grow epitaxial thin films of $YbRh_2Si_2$ on (terahertz transparent) Ge substrates (Fig. 1A), we used a specially equipped MBE system (*18*). The epitaxial growth of phase-pure $YbRh_2Si_2$ was confirmed with x-ray diffraction (Fig. 1B) (*18*), and the high quality of the film and the film-substrate interface were revealed with high-resolution transmission electron microscopy (Fig. 1, C and D) (*18*). The temperature dependence of the (quasi) dc electrical resistivity $\rho(T)$ of these films (*18*) is similar to that of bulk single crystals (Fig. 2) (*15*, *19*). $\rho(T)$ displays strange-metal behavior, $\rho = \rho_0 + A'T^{\alpha}$ (Fig. 2B), where A' is a constant, with an exponent α that strongly deviates from the Fermi liquid value $\alpha = 2$ and tends to $\alpha = 1$ in the low-temperature limit (fig. S1).

The frequency dependence of the real part of the complex optical conductivity, $\mathrm{Re}(\sigma)$, measured at temperatures between 1.4 and 250 K and frequencies between 0.25 and 2.6 THz, is shown in Fig. 3A [the imaginary part, $\mathrm{Im}(\sigma)$, is shown in fig. S2]. The dc electrical conductivity $\sigma = 1/\rho$ values, plotted as symbols at $\omega = 0$, are compatible with the extrapolation of the finite frequency results to zero frequency. Both $\mathrm{Re}(\sigma)$ and $\mathrm{Im}(\sigma)$ are flat and featureless at temperatures above ~80 K (indicating strong incoherent scattering of charges) but develop sizable temperature and frequency dependence at lower temperatures, with spectral weight of $\mathrm{Re}(\sigma)$ being transferred to low frequencies. The increasingly sharp and pronounced resonance of $\mathrm{Re}(\sigma)$, with non-Lorentzian shape (non-Drude behavior) (fig. S3), may in clean samples be associated with non–Fermi liquid behavior. These results confirm deviations from simple Drude behavior seen earlier in optical reflectivity measurements in the far-infrared range on bulk $YbRh_2Si_2$ single crystals (*20*).

To explore dynamical scaling, we analyzed the frequency-dependent intrinsic optical conductivity $\sigma_{in}(\omega)$ by subtracting a residual resistivity because of impurity scattering; this subtraction is motivated by analogy to the Matthiessen's law used for the dc resistivity (*18*). We plot $\mathrm{Re}[\sigma_{in}(\omega)] \cdot T^{\alpha}$ as a function of $\hbar\omega/(k_B T)$, where $\hbar$ is the Planck constant divided by 2π and k_B is the Boltzmann constant, for temperatures ($T \leq 15$ K) well below the material's Kondo temperature $T_K = 24$ K (Fig. 3B) (*15*) and frequencies below 2 THz. For $\alpha \approx 1$, all curves collapse, demonstrating ω/T scaling of $\mathrm{Re}[\sigma_{in}(\omega)]$.

How can the optical conductivity, which probes charge fluctuations, show critical ω/T scaling at an antiferromagnetic quantum critical point where a priori only spin fluctuations are expected—and indeed observed (*21*–*23*)—to be critical? A natural way for this to happen is to have a critical form of the Kondo entanglement between the local moments and the conduction electrons (*24*–*26*), as illustrated in Fig. 4. Across the quantum critical point, the conduction electrons go from being (asymptotically) decoupled from the local moments (Fig. 4, bottom left box) to being entangled with them (Fig. 4, bottom right box). Correspondingly, the elementary excitations change from separate charge (single conduction electrons or holes) and spin excitations (Fig. 4, top left box) to the heavy quasiparticles (Fig. 4, top right box) that are hybrids of the slow composite fermions (Fig. 4, large tadpole) and the bare conduction electrons (Fig. 4, small tadpole); the single-electron excitations capture the continuous onset of the Kondo entanglement at the quantum critical point and are part of the critical degrees of freedom. Thus, optical conductivity, which probes the charge current of the elementary excitations, manifests the singular fluctuations of the quantum critical point. Within the Landau description of a metallic antiferromagnetic quantum critical point (*1*, *2*),

[1]Institute of Solid State Physics, Technischen Universität (TU) Wien, Wiedner Hauptstraße 8-10, 1040 Vienna, Austria. [2]Department of Electrical and Computer Engineering, 6100 Main Street, Rice University, Houston, TX 77005, USA. [3]Institute of Solid State Electronics, TU Wien, Nanocenter Campus Gußhaus, Gußhausstraße 25-25a, Gebäude CH, 1040 Vienna, Austria. [4]Institute of Chemical Technologies and Analytics, TU Wien, Getreidemarkt 9, 1060 Vienna, Austria. [5]Department of Chemistry, 6100 Main Street, Rice University, Houston, TX 77005, USA. [6]Department of Materials Science and Nanoengineering, 6100 Main Street, Rice University, Houston, TX 77005, USA. [7]Center for Micro- and Nanostructures, TU Wien, Nanocenter Campus Gußhaus, Gußhausstraße 25-25a, Gebäude CH, 1040 Vienna, Austria. [8]Department of Physics and Astronomy, Center for Quantum Materials, 6100 Main Street, Rice University, Houston, TX 77005, USA.
*These authors contributed equally to this work. †Present address: Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA; and Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA 91125, USA. ‡Present address: State University of New York–University at Buffalo, Jacobs School of Medicine and Biomedical Sciences, 955 Main Street, Buffalo, NY 14203, USA. §Present address: Kavli Institute at Cornell for Nanoscale Science, Cornell University, Ithaca, NY 14853, USA. ¶Present address: Renewable and Sustainable Energy Institute, University of Colorado Boulder, Boulder, CO 80309, USA. #Present address: Central European Institute of Technology, Brno University of Technology, Purkyňova 123, Brno, 612 00, Czech Republic. **Present address: Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK; and Department of Materials Science and Metallurgy, University of Cambridge, 27 Charles Babbage Road, Cambridge CB3 0FS, UK.
††Corresponding author. Email: paschen@ifp.tuwien.ac.at

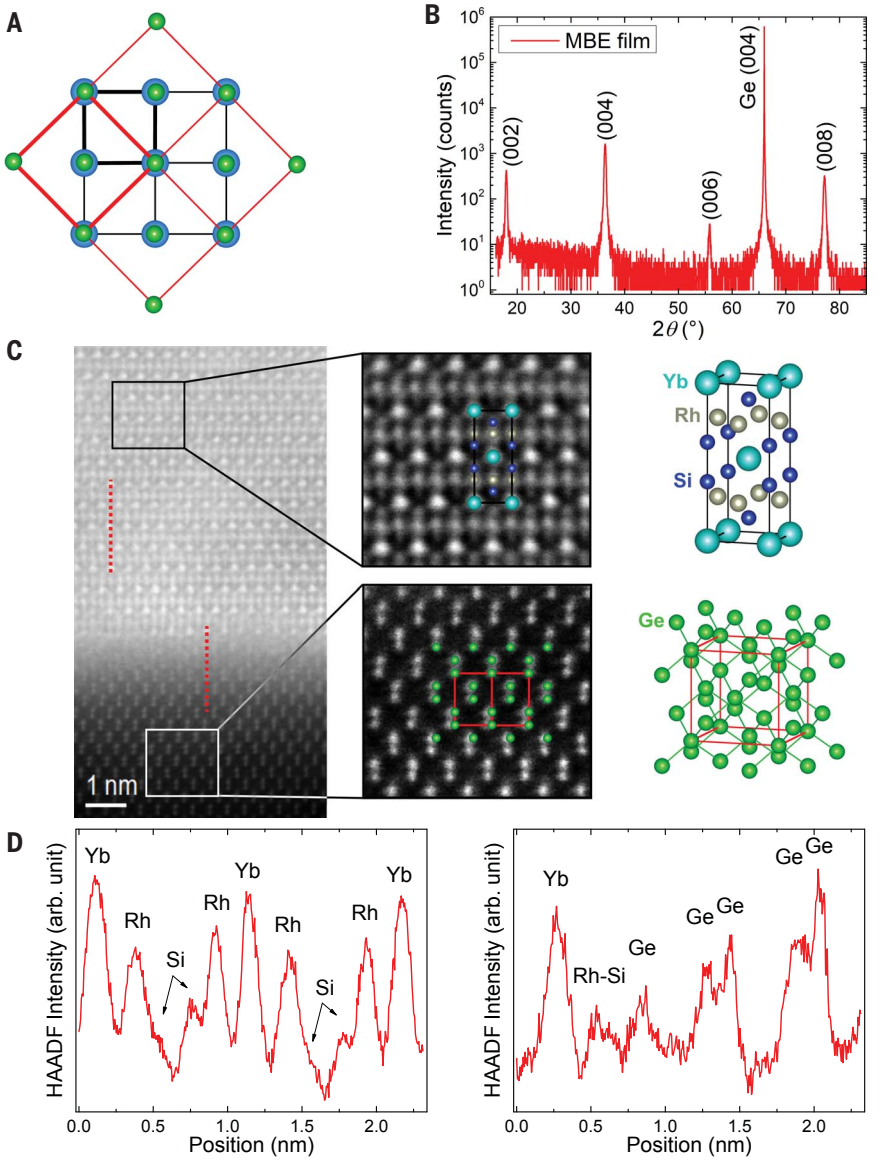


Fig. 1. $YbRh_2Si_2$ thin films grown by means of MBE. (**A**) Visualization of the lattice matching between $YbRh_2Si_2$ (blue circles and black lines) and Ge (green circles and red lines), with the crystallographic *c* directions pointing out of the plane. For the Yb atoms to associate with the Ge atoms, the respective unit cells (thick lines) [(C), right] are rotated by 45° with respect to each other around the *c* direction. (**B**) High-resolution x-ray diffraction pattern, with all peaks identified as due to the (about 40 nm thick) film or the Ge substrate, confirming that the film is phase-pure $YbRh_2Si_2$. (**C**) Atomic-resolution high-angle annular dark field scanning transmission electron microscopy (HAADF-STEM) image of the interface between film (top left) and substrate (bottom left), representative enlarged views with simulated overlays (center), and the corresponding unit cells (right). (**D**) Intensity profiles along the red dotted lines in (C). The left and right panels correspond to the top red dotted line within the film and the bottom red dotted line across the interface, respectively.

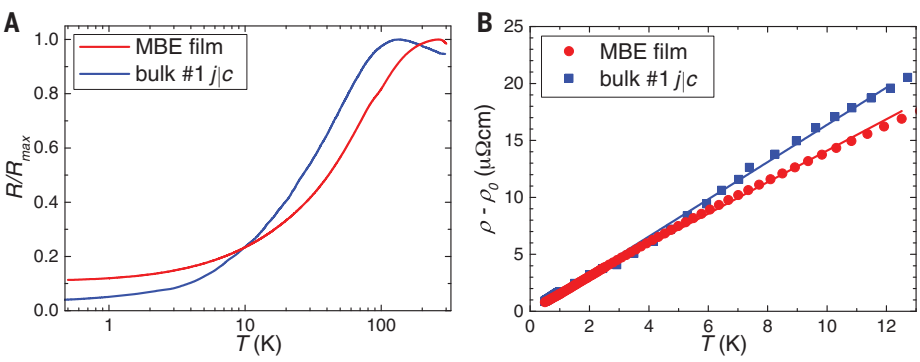


Fig. 2. Electrical resistivity of MBE-grown $YbRh_2Si_2$. (**A**) Normalized resistance of an MBE-grown $YbRh_2Si_2$ film and a bulk single crystal with current *j* within the tetragonal *aa* plane (*15*) for comparison. The film was measured by using the van der Pauw technique. (**B**) Corresponding low-temperature resistivities, with the residual resistivities [ρ_0 = 11.6 and 2.45 microhm cm for the MBE and bulk (*15*) samples, respectively, determined by linear-in-*T* fits to the data below 1 K] subtracted, displaying non–Fermi liquid behavior (lines represent $\rho - \rho_0 = A'T^{\alpha}$ fits with a constant exponent α to the data below 12 K; the temperature dependence of α is provided in fig. S1).

the slow long-wavelength fluctuations of the order parameter alone describe a Gaussian fixed point, where ω/T scaling is violated. The incorporation of the single-electron excitations in the quantum critical spectrum not only makes charge fluctuations part of the quantum criticality but also turns the fixed point into an interacting one (*24*), leading to ω/T scaling.

Dynamical scaling of the optical conductivity in the region of *T*-linear resistivity has also been analyzed in an optimally doped Bi-2212 cuprate (*27*). There, different scaling functions are needed in different ω/T ranges, leaving open the question of how the fluctuations of the charge carriers connect with the robust linear-in-temperature resistivity of the cuprate superconductors. By contrast, in the present study of $YbRh_2Si_2$, a single ω/T scaling form

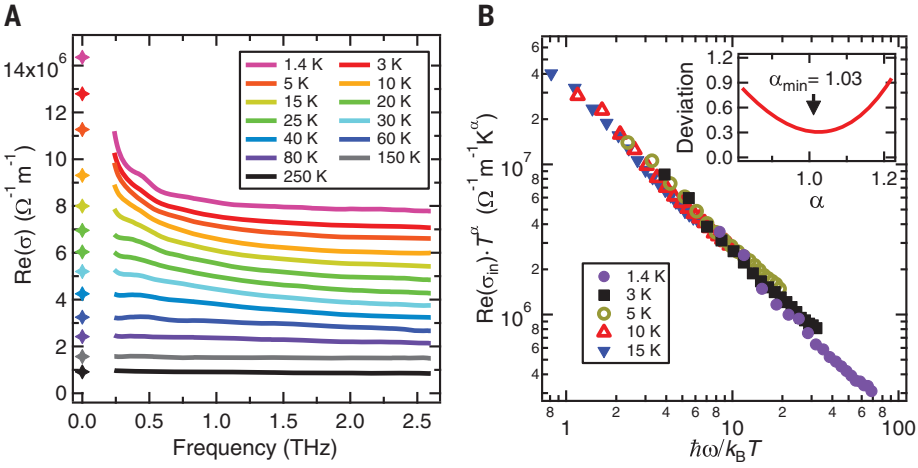


Fig. 3. Terahertz time-domain transmission spectroscopy of MBE-grown $YbRh_2Si_2$. (**A**) Real part of optical conductivity Re(σ) versus frequency at different temperatures (bottom to top: 250, 150, 80, 60, 40, 30, 25, 20, 15, 10, 5, 3, and 1.4 K), with corresponding dc values marked as zero-frequency points. Curves below 250 K (and the respective dc values) are successively offset by 6×10^5 ohm^{-1} m^{-1} for clarity. (**B**) ω/T scaling, with a critical exponent of $\alpha \approx 1$, revealed with $\mathrm{Re}[\sigma_{in}(\omega)] \cdot T^{\alpha}$ isotherms plotted versus $\hbar\omega/(k_B T)$ collapsing onto a single curve for temperatures $T \leq 15$ K and frequencies below 2 THz. (Inset) Normalized deviation between the different isotherms as a function of α, revealing best scaling for $\alpha = 1.03$.

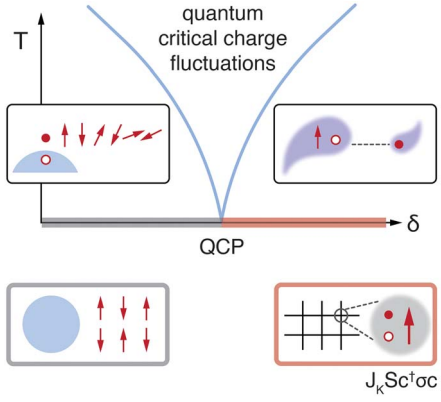


Fig. 4. Illustration of quantum-critical charge fluctuations emerging from Kondo disentanglement. Tuning a heavy fermion metal with a nonthermal parameter δ, which microscopically corresponds to the ratio of Kondo to RKKY coupling, from an antiferromagnetic ground state with local moment order (bottom left box; blue circle and red arrows indicate Fermi sphere and local moments, respectively) to a Kondo entangled paramagnet (bottom right box; the antiferromagnetic Kondo exchange J_K favors the formation of a Kondo singlet between the local moment S, represented as an arrow, and the spin of the conduction electrons $c^{\dagger}\sigma c$—the particle-hole excitation of the Fermi sea in the spin-triplet channel) creates distinct single-particle excitations (top boxes) and, in turn, quantum-critical charge fluctuations within the quantum-critical fan.

is uncovered in its strange metal regime. It is important to explore the dynamical scaling of the optical conductivity in other materials classes with strange-metal behavior; one can then assess whether the charge carrier dynamics emerging from a localization-delocalization quantum critical point, as proposed here, is a universal mechanism of strange-metal behavior. This scaling form also provides an intriguing link to the quantum scaling of metal-insulator transitions, both in Mott-Hubbard (*28*–*30*) and in disordered systems (*31*).

Our results demonstrate that charge carriers are a central ingredient of the singular physics at the border of antiferromagnetic order, providing direct evidence for the beyond-Landau nature of metallic quantum criticality. Our findings also delineate the role of electronic localization transitions in strange-metal phenomena, which are relevant to a variety of strongly correlated materials (*32*) and beyond (*33*).

REFERENCES AND NOTES

1. J. A. Hertz, *Phys. Rev. B Condens. Matter* **14**, 1165–1184 (1976).
2. A. J. Millis, *Phys. Rev. B Condens. Matter* **48**, 7183–7196 (1993).
3. D. Bitko, T. F. Rosenbaum, G. Aeppli, *Phys. Rev. Lett.* **77**, 940–943 (1996).
4. P. Merchant *et al.*, *Nat. Phys.* **10**, 373–379 (2014).
5. P. Giraldo-Gallo *et al.*, *Science* **361**, 479–481 (2018).
6. J. A. N. Bruin, H. Sakai, R. S. Perry, A. P. Mackenzie, *Science* **339**, 804–807 (2013).
7. S. Paschen *et al.*, *Nature* **432**, 881–885 (2004).
8. S. Friedemann *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 14547–14551 (2010).
9. J. Custers *et al.*, *Nat. Mater.* **11**, 189–194 (2012).
10. Y. Luo *et al.*, *Nat. Mater.* **13**, 777–781 (2014).
11. F. F. Balakirev *et al.*, *Nature* **424**, 912–915 (2003).
12. H. Oike, K. Miyagawa, H. Taniguchi, K. Kanoda, *Phys. Rev. Lett.* **114**, 067002 (2015).
13. S. Badoux *et al.*, *Nature* **531**, 210–214 (2016).
14. Y. Cao *et al.*, *Nature* **556**, 43–50 (2018).
15. O. Trovarelli *et al.*, *Phys. Rev. Lett.* **85**, 626–629 (2000).
16. P. Gegenwart *et al.*, *Phys. Rev. Lett.* **89**, 056402 (2002).
17. J. Orenstein *et al.*, *Phys. Rev. B Condens. Matter* **42**, 6342–6362 (1990).
18. Materials and methods are available as supplementary materials.
19. T. Westerkamp, thesis, TU Dresden, Germany (2009).
20. S. Kimura *et al.*, *Phys. Rev. B Condens. Matter* **74**, 132408 (2006).
21. M. C. Aronson *et al.*, *Phys. Rev. Lett.* **75**, 725–728 (1995).
22. A. Schröder *et al.*, *Nature* **407**, 351–355 (2000).
23. C. Stock *et al.*, *Phys. Rev. Lett.* **109**, 127201 (2012).
24. Q. Si, S. Rabello, K. Ingersent, J. L. Smith, *Nature* **413**, 804–808 (2001).
25. P. Coleman, C. Pépin, Q. Si, R. Ramazashvili, *J. Phys. Condens. Matter* **13**, R723–R738 (2001).
26. T. Senthil, M. Vojta, S. Sachdev, *Phys. Rev. B Condens. Matter Mater. Phys.* **69**, 035111 (2004).
27. D. van der Marel *et al.*, *Nature* **425**, 271–274 (2003).
28. T. Senthil, *Phys. Rev. B Condens. Matter* **78**, 035103 (2008).
29. H. Terletska, J. Vučičević, D. Tanasković, V. Dobrosavljević, *Phys. Rev. Lett.* **107**, 026401 (2011).
30. T. Furukawa, K. Miyagawa, H. Taniguchi, R. Kato, K. Kanoda, *Nat. Phys.* **11**, 221–224 (2015).
31. H. Lee, J. P. Carini, D. V. Baxter, W. Henderson, G. Grüner, *Science* **287**, 633–636 (2000).
32. B. Keimer, J. E. Moore, *Nat. Phys.* **13**, 1045–1055 (2017).
33. P. T. Brown *et al.*, *Science* **363**, 379–382 (2019).
34. L. Prochaska, X. Li, D. C. MacFarland, A. M. Andrews, M. Bonta, E. F. Bianco, S. Yazdi, W. Schrenk, H. Detz, A. Limbeck, Q. Si, E. Ringe, G. Strasser, J. Kono, S. Paschen, Data for the publication "Singular charge fluctuations at a magnetic quantum critical point." Zenodo (2019); doi:10.5281/zenodo.3541572.

ACKNOWLEDGMENTS

We thank P. Gegenwart, Y.-B. Kim, H. von Löhneysen, S. Nakatsuji, H.-C. Nägerl, and A. Prokofiev for useful discussions. **Funding:** Financial support for this work was provided by the European Research Council (ERC Advanced Grant 227378), the U.S. Army Research Office (ARO W911NF-14-1-0496), the Austrian Science Fund (FWF W1243, P29279-N27, and P29296-N27), and the European Union's Horizon 2020 research and innovation program (grant agreement 824109-EMP). X.L. and J.K. acknowledge financial support from the National Science Foundation (NSF MRSEC DMR-1720595) and the ARO (W911NF-17-1-0259). Q.S. acknowledges financial support from the NSF (DMR-1920740), the Robert A. Welch Foundation (C-1411), and the ARO (W911NF-14-1-0525) and the hospitality of the University of California at Berkeley, the Aspen Center for Physics (NSF grant PHY-1607611), and the Los Alamos National Laboratory (through a Ulam Scholarship from the Center for Nonlinear Studies). This work has also been supported by an InterDisciplinary Excellence Award (IDEA) from Rice University (Q.S., E.R., J.K., and S.P.). **Author contributions:** S.P. designed and led the research. L.P., D.C.M., A.M.A., W.S., H.D., and G.S. performed the MBE growth. M.B. and A.L. performed the analytical characterization. E.F.B., S.Y., and E.R. performed the TEM investigation. X.L. and J.K. performed the terahertz spectroscopy. Q.S. contributed to the understanding of the results. L.P., X.L., D.C.M., Q.S., and S.P. wrote the manuscript, with contributions from all other authors. **Competing interests:** The authors have no competing interests. **Data and materials availability:** All data presented in this paper are deposited in Zenodo (*34*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/285/suppl/DC1
Materials and Methods
Figs. S1 to S6
References (*35*–*53*)

6 August 2018; resubmitted 7 September 2019
Accepted 5 December 2019
10.1126/science.aag1595

OPTICS

Subwavelength dielectric resonators for nonlinear nanophotonics

Kirill Koshelev[1,2], Sergey Kruk[1], Elizaveta Melik-Gaykazyan[1,3], Jae-Hyuck Choi[4], Andrey Bogdanov[2], Hong-Gyu Park[4]*, Yuri Kivshar[1,2]*

Subwavelength optical resonators made of high-index dielectric materials provide efficient ways to manipulate light at the nanoscale through mode interferences and enhancement of both electric and magnetic fields. Such Mie-resonant dielectric structures have low absorption, and their functionalities are limited predominantly by radiative losses. We implement a new physical mechanism for suppressing radiative losses of individual nanoscale resonators to engineer special modes with high quality factors: optical bound states in the continuum (BICs). We demonstrate that an individual subwavelength dielectric resonator hosting a BIC mode can boost nonlinear effects increasing second-harmonic generation efficiency. Our work suggests a route to use subwavelength high-index dielectric resonators for a strong enhancement of light–matter interactions with applications to nonlinear optics, nanoscale lasers, quantum photonics, and sensors.

High-index resonant dielectric nanostructures emerged recently as a new platform for nano-optics and photonics to complement plasmonic structures in a range of functionalities (*1*, *2*). All-dielectric nanoresonators benefit from low material losses and allow the engineering of artificial magnetic responses. Progress in all-dielectric nanophotonics led to the development of efficient flat-optics devices that reached and even outperformed the capabilities of conventional bulk components (*3*). These advances motivated further diversification of applications of dielectric nanostructures (*4*), especially toward nonlinear optics (*5*–*7*). Efficiencies of nonlinear optical processes in all-dielectric nanostructures have exceeded by several orders of magnitude the efficiencies demonstrated in metallic nanoparticles with plasmonic resonances (*8*, *9*).

One of the main limiting factors for high efficiencies of all-dielectric nanostructures as functional devices is the quality (Q) factor of their resonant modes. Traditionally, response of dielectric nanoparticles is governed by low-order geometrical resonances, resulting in low Q factors. An elegant solution for Q-factor control and its increase is provided by the physics of bound states in the continuum (BICs). BICs were first proposed in quantum mechanics as localized electron waves with the energies embedded within the continuous spectrum of propagating waves (*10*). Recently, BICs have attracted considerable attention in photonics (*11*, *12*). Mathematical bound states have infinitely large Q factors and vanishing resonant linewidth. In practice, BICs are limited by a finite sample size, material absorption, and structural imperfections (*13*), but they manifest themselves as resonant states with large Q factors, also known as quasi-BICs or supercavity modes. Until now, optical BICs have been observed only for extended systems (*12*, *14*–*16*) and used for various applications including lasing (*17*) and sensing (*18*). For individual isolated dielectric resonators, genuine nonradiative states require extreme material parameters diverging toward infinity or zero (*19*, *20*). In realistic individual resonators, there are an infinite number of possible paths for radiation to escape (*12*), which limits the Q factor substantially. However, the concept of quasi-BICs allows us to come close to reaching nonradiative states for individual dielectric resonators (*21*–*23*). The modes forming a quasi-BIC belong to the same resonator, which allows the system footprint to be kept very small. Except for specific composite structures (*24*), such a small footprint is challenging to achieve for resonators relying on alternative mechanisms of localization, including whispering gallery mode resonators and cavities in photonic bandgap structures.

Here, we studied individual dielectric nanoresonators hosting a quasi-BIC resonance at telecommunication wavelengths and demonstrated its capability for second-harmonic generation (SHG). Our subwavelength resonator exploits mutual interference of several Mie modes, which results in a quasi-BIC regime. We designed a 635-nm-tall nonlinear nanoresonator of cylindrical shape made of AlGaAs (aluminum gallium arsenide) placed on an engineered three-layer substrate (SiO_2/ITO/SiO_2) (Fig. 1A). For a cylindrical particle, Mie resonances are classified with an azimuthal order and can be loosely sorted into two groups distinguished by the number of oscillations in the radial and axial directions [see part 1 of the supplementary text (*25*)]. We selected a pair of modes from different groups with uniform azimuthal field distribution (Fig. 1B), both of which demonstrated a magnetic dipolar behavior [see parts 1 and 2 of the supplementary text (*25*)]. By changing the resonator's diameter, the spectral mismatch of dipolar modes can be decreased, which induces their strong coupling in the parametric space and produces the characteristic avoided resonance crossing of frequency curves (Fig. 1C). In the strong coupling regime, the modes are hybrid with a combination of radial or axial oscillations and thus do not belong to any of the defined groups [see part 1 of the supplementary text (*25*)]. Open boundaries of the nanoresonator enable constructive and destructive mode interference in the far field (*26*), which results in modification of the mode Q factors because of their identical dipolar nature (Fig. 1D). The quasi-BIC regime with suppressed dipolar radiation (Fig. 1E) and thus an increased Q factor is reached for a particle of a specific diameter of ~930 nm.

We further compensated for the decrease of the Q factor induced by energy leakage into the substrate (*27*) by adding a layer of ITO (indium tin oxide) exhibiting an epsilon-near-zero transition acting as a conductor above a 1200-nm wavelength (e.g., at the quasi-BIC wavelength) and as an insulator below this wavelength [e.g., at the second harmonic (SH) wavelength]. The ITO layer is separated from the resonator by a SiO_2 spacer. The thickness of the SiO_2 spacer layer provides control over the phase of reflection, further enhancing the destructive interference of the two magnetic dipoles in the far field and thus increasing the Q factor (Fig. 1F). For the optimal spacer thickness between 300 and 400 nm, the Q factor reaches the maximal predicted value of 235.

We fabricated a set of individual AlGaAs nanoresonators with diameters varying from 890 to 980 nm from epitaxially grown AlGaAs (crystal axes orientation [100], 20% Al) by means of electron-beam lithography and a dry-etching process. The nanoparticles were subsequently transferred to a substrate made of a commercial film of 300-nm ITO on glass with an added SiO_2 spacer 350-nm thick [see materials and methods and part 7 of the supplementary text (*25*)]. We measured scattering spectra from individual nanoparticles with a laser tunable within the wavelength range of 1500 to 1700 nm. To maximize light coupling to the quasi-BIC mode, we illuminated each nanoresonator with a tightly focused, azimuthally polarized light [see the materials and methods and part 8 of the supplementary text (*25*)]. The scattering spectra are evaluated as the difference between the bare substrate

[1]Nonlinear Physics Center, Australian National University, Canberra ACT 2601, Australia. [2]Department of Physics and Engineering, ITMO University, St. Petersburg 197101, Russia. [3]Faculty of Physics, Lomonosov Moscow State University, Moscow 119991, Russia. [4]Department of Physics, Korea University, Seoul 02841, Republic of Korea.
*Corresponding author. Email: hgpark@korea.ac.kr (H.-G.P.); ysk@internode.on.net (Y.K.)

reflectivity and the normalized measured backward scattering of the nanoresonator. We observed a symmetric peak with the extracted Q factor of 188 ± 5 for the particle diameter of ~930 nm, which corresponds to the quasi-BIC condition [see Fig. 1G and the materials and methods (*25*) for details on the Q-factor extraction procedure]. We further measured the dependence of the Q factor on the nanoresonator diameter (dots in Fig. 1D), which showed good agreement with numerical simulations.

Next, we exploited the designed quasi-BIC resonator as a nonlinear nanoantenna for SH generation (Fig. 2A). At the SH wavelength, the nanoresonator supports a high-order Mie mode with a Q factor of 65 [see part 1 of the supplementary text (*25*)]. For SH wavelengths, the material properties of ITO are similar to glass, so the spacer and ITO thickness are inessential. To increase the nonlinear conversion efficiency, we developed the consistent theory of SHG for nanoscale resonators using the eigenmode expansion method [see parts 4 and 5 of the supplementary text (*25*)], which goes beyond the phase-matching approach used for nonlinear optics of macroscopic structures (*28*).

The optical response of designed nonlinear nanoantenna is driven by the quasi-BIC with complex frequency $\omega_1 - i\gamma_1$ and the SH Mie mode with frequency $\omega_2 - i\gamma_2$. The total SH power radiated by the nanoresonator [see part 5 of the supplementary text (*25*)] is:

$$P^{2\omega} = \alpha\ \kappa_2\, Q_2\, L_2\, \kappa_{12}\, [Q_1\, L_1\, \kappa_1\, P^{\omega}]^2 \qquad (1)$$

This expression allows a step-by-step explanation of the SHG process (Fig. 2B). The incident power P^{ω} is coupled to the quasi-BIC depending on the spatial overlap κ_1 between the pump and the mode. The coupled power is resonantly enhanced depending on the quasi-BIC Q factor Q_1 and damped by the spectral overlap factor $L_1(\omega) = \gamma_1^2 / [(\omega - \omega_1)^2 + \gamma_1^2]$, which is the unity at the resonance. The efficiency of upconversion of the total accumulated power is determined by the cross-coupling coefficient κ_{12}, which depends on the symmetry of the nonlinear susceptibility tensor of AlGaAs and the spatial overlap between the generated nonlinear polarization current and SH mode [see part 5 of the supplementary text (*25*)]. The converted SH power is increased by a high Q factor of the SH mode but at the same time is decreased because of the spectral mismatch with the quasi-BIC, $L_2(2\omega_1) = \gamma_2^2 / [(2\omega_1 - \omega_2)^2 + \gamma_2^2]$. The out-coupling factor $\kappa_2(2\omega)$ determines a fraction of the radiated SH power and is the unity in

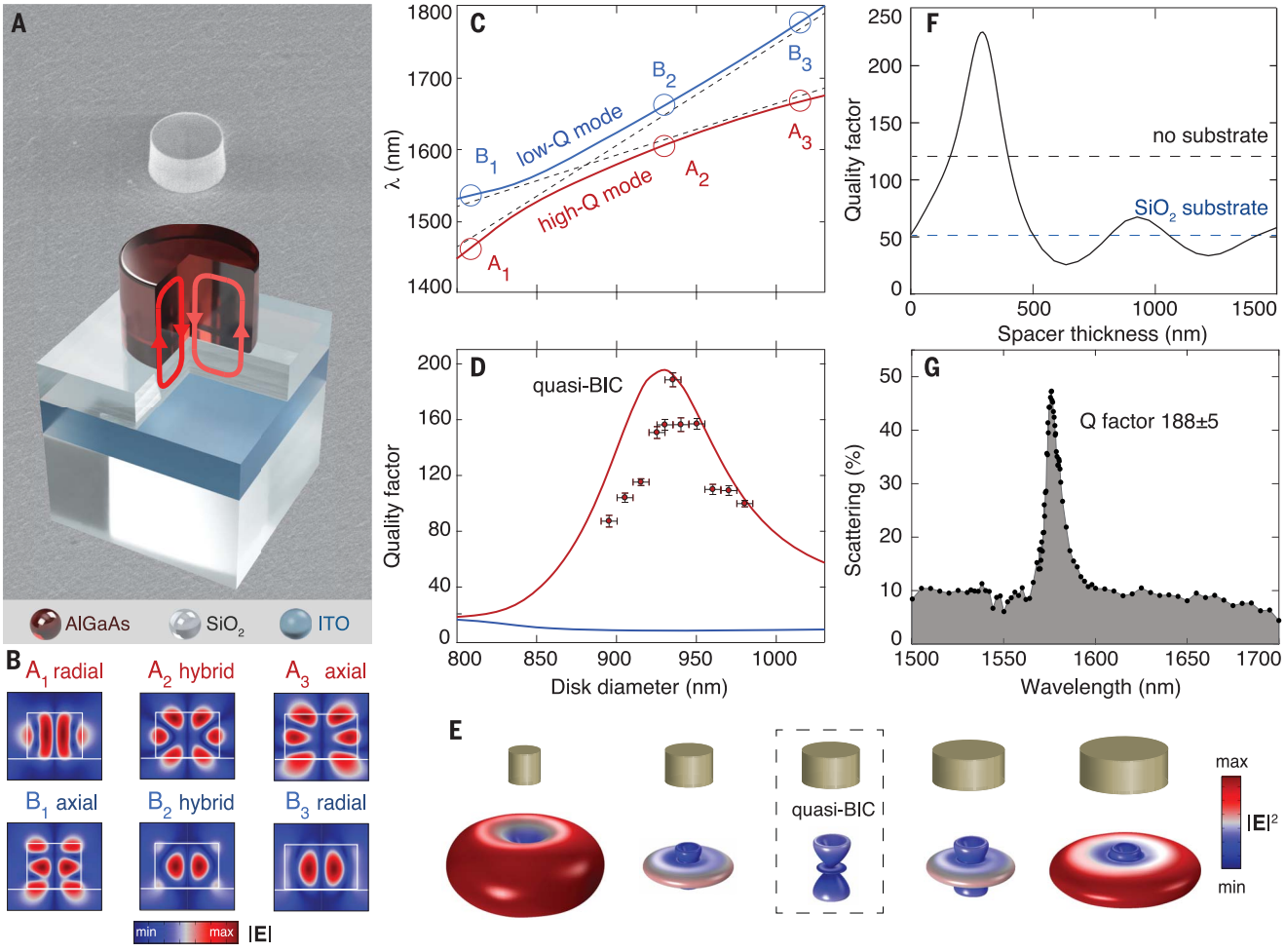


Fig. 1. Optical quasi-BIC mode in an individual dielectric nanoresonator. (**A**) Scanning electron micrograph (top) and schematic (bottom) of an individual dielectric nanoresonator. (**B**) Simulated near-field patterns of the two modes for different diameters. (**C**) Calculated mode wavelengths versus resonator diameter. (**D**) Calculated (lines) and measured (dots) Q factors of modes versus resonator diameter. Calculations in (C) and (D) are done for a 350-nm SiO_2 spacer. (**E**) Simulated far-field patterns of the high-Q mode for disks of different diameters shown schematically. For calculations, $|\mathbf{E}|^2$ is normalized to the full mode energy. (**F**) Calculated Q factor of the quasi-BIC versus SiO_2 spacer thickness compared with the Q factors of a nanoresonator in air and on a bulk SiO_2 substrate (dashed lines). (**G**) Measured scattering spectrum and retrieved Q factor of the observed resonance for a disk with a diameter of ~930 nm.

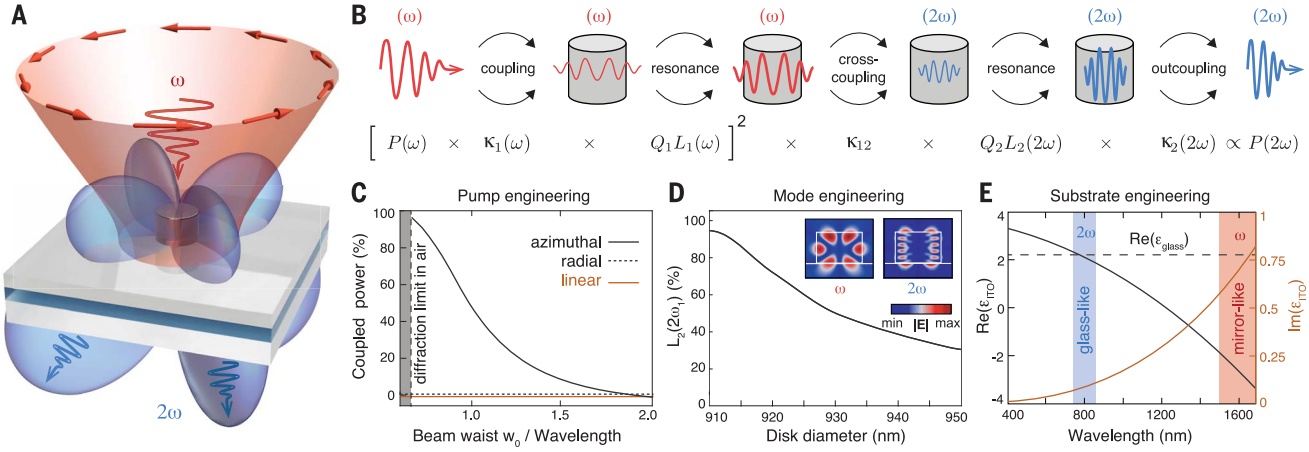


Fig. 2. Second-harmonic generation with a dielectric nanoantenna. (**A**) Diagram of the SHG in a nanoresonator under azimuthally polarized vector beam excitation. (**B**) Schematic of the SHG process in a nonlinear dielectric nanoantenna. Each term of the formula describes one step of the process. (**C**) Percentage of pump power coupled to the quasi-BIC for different polarizations of pump depending on the ratio between the beam waist radius w_0 and the pump wavelength. The calculation is done for a free-standing nanoresonator in air. The diffraction limit is 0.61. (**D**) Spectral overlap $L_2(2\omega_1)$ between the high-Q mode at the pump frequency and the high-order Mie mode at the SH frequency versus the disk diameter. The inset shows the near-field profiles of both modes. (**E**) Experimental ellipsometry data for the permittivity of the ITO layer. Wavelength ranges of the excitation and collection are marked with red and blue shading, respectively.

the vicinity of ω_2. The exact expressions for coupling coefficients κ_1, κ_{12}, and κ_2 and the constant α are given in part 5 of the supplementary text (*25*). Note that the effective mode volume does not appear in Eq. 1 because κ_{12} takes into account the explicit spatial distributions of the electric field of the modes.

With this theoretical analysis, we can specify the optimal conditions to maximize the SHG efficiency from an individual dielectric nanoresonator. First, the spatial profile of the pump must be structured to match the distribution of the excited mode; therefore, we used the cylindrical vector beam with azimuthal polarization. We estimated κ_1 as 33% for the experimental conditions using a model of a free-standing resonator in air (Fig. 2C) [see parts 5 and 10 of the supplementary text (*25*)]. Next, the optimal structure must be resonant simultaneously at pump and SH wavelengths (*9*). Maximization of Q_1 is critical compared with maximization of Q_2 because of the quadratic over linear dependence of $P^{2\omega}$ [see Eq. (1) and part 6 of the supplementary text (*25*)]. For the designed nanoresonator with a diameter of ~930 nm, the factor of spectral overlap reaches 50% (Fig. 2D). Finally, the collection efficiency must be increased, which can be achieved by engineering the substrate properties. The epsilon-near-zero transition of ITO makes it effectively "invisible" to the SH radiation, allowing it to propagate in both the forward and backward directions (Fig. 2E).

To perform systematic experimental analysis of the SHG enhancement in quasi-BIC resonators, we excited the fabricated set of nanoparticles with laser pulses of 2-ps duration in the wavelength range from 1500 to 1700 nm [see the materials and methods and part 9 of the supplementary text (*25*)]. Figure 3, A to D, shows the maps of the SHG intensity versus the pump wavelength and resonator diameter for the nanoresonators pumped by the azimuthal, radial, and linearly polarized beams, respectively. The experimental data reveal a sharp enhancement of the nonlinear signal in the quasi-BIC regime selectively for the azimuthally polarized pump. We measured directionality diagrams of the SH signal in the backward and forward directions within the numerical apertures of a pair of confocal objective lenses [see Fig. 3, E and F, and part 12 of the supplementary text (*25*)]. The diagram in the backward direction features distinct maxima in four directions that are qualitatively similar to the theoretical SHG directionality shown in Fig. 2A and the far-field pattern of the mode excited at the SH wavelength [see part 1 of the supplementary text (*25*)].

Figure 4, A and B, shows a wavelength cut (at the quasi-BIC diameter of ~930 nm) and a size cut (at the quasi-BIC wavelength of 1570 nm) of the measured 2D SHG maps (see Fig. 3, B to D). Both plots demonstrate that the observed SH intensity for the azimuthal pump surpasses the SH intensity for the other polarizations by several orders of magnitude, which confirms high spatial selectivity of the quasi-BIC [see also part 3 of the supplementary text (*25*)]. With these experiments, we reached beyond the predictions of the theoretical model (see Fig. 2C) and measured an observable SH signal for radial and linear polarizations caused by off-resonant excitation of other nanoparticle modes (Fig. 4B). However, this signal remains several orders of magnitude lower compared with azimuthal polarization. We further experimentally measured the SHG conversion efficiency. The numerical analysis of quasi-BICs in a nonlinear nanoresonator [see Fig. 2 and (*23*)] does not account for the trade-off between pulse duration and laser damage threshold. The high Q factor of the quasi-BIC requires relatively long pulses to pump the mode effectively. At the same time, a peak power of longer pulses becomes limited by the material laser damage threshold. From this point of view, theoretical or numerical analysis does not answer the question of whether a nanoresonator made of common dielectric materials can indeed function as an efficient nonlinear nanoantenna.

We conducted an experimental verification of this by detecting the peak pump power P_p^{ω} incident onto the sample and the peak SH power $P_p^{2\omega}$ captured by the two objective lenses in the forward and backward directions (Fig. 4C). The directly measured conversion efficiency $P_p^{2\omega}/(P_p^{\omega})^2$ was 1.3×10^{-6} W^{-1} [see part 11 of the supplementary text (*25*)]. The observed SHG efficiency at the quasi-BIC was more than two orders of magnitude higher than that demonstrated with earlier implementations using other approaches (*5*–*7*, *9*). We further estimate the total SHG efficiency as 4.8×10^{-5} W^{-1} using the common approach by taking into account only the coupled part of P_p^{ω}, theoretically estimated as 33%, and the total SH power, estimated using the calculated collection efficiency of 24%. A detailed list of the

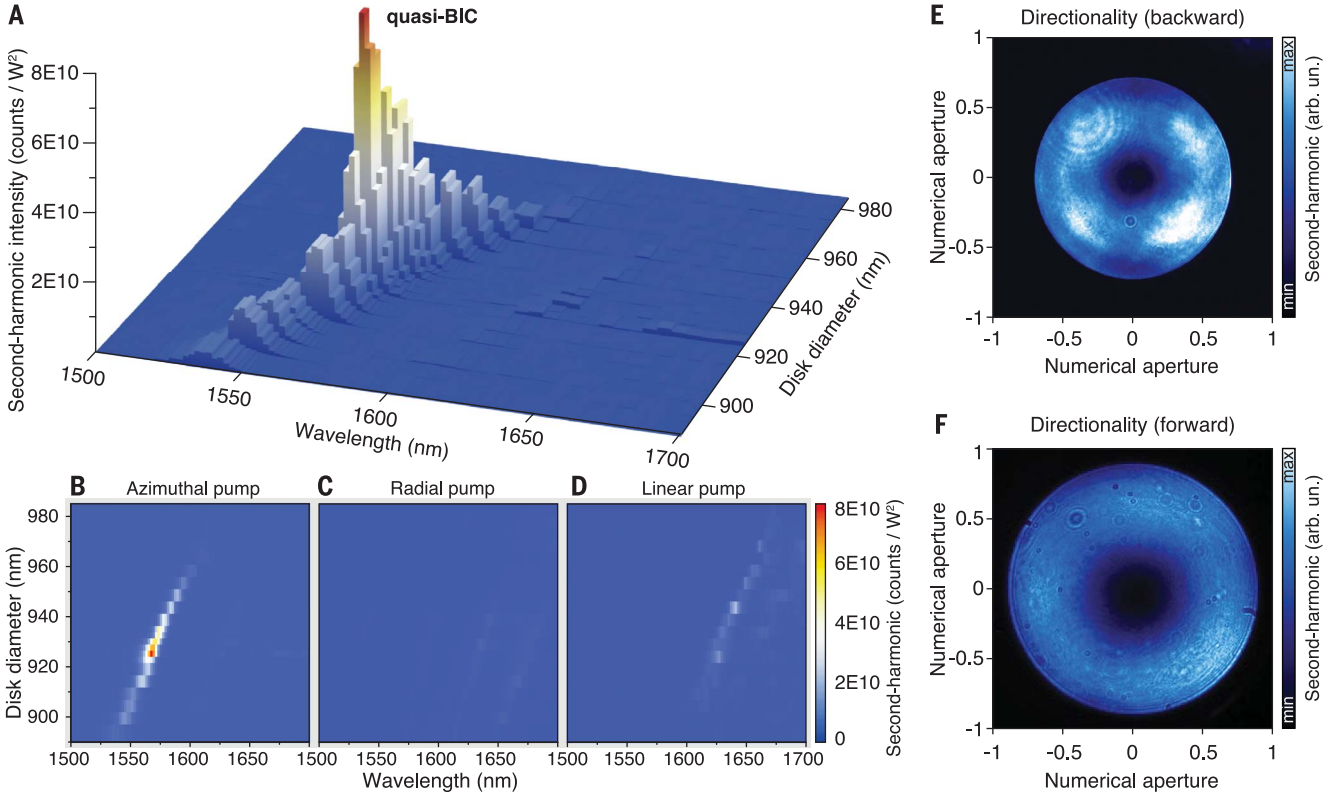


Fig. 3. Experimental characterization of the SHG enhancement. (**A**) 3D map of SH intensity measured as a function of the pump wavelength and particle diameter for an azimuthally polarized beam. The SH intensity is normalized on the square of the pump power. (**B** to **D**) Top views of the maps of SHG with the azimuthal, radial, and linear pump, respectively. (**E** and **F**) Experimentally measured directionality diagrams of SHG for a nanoresonator with a diameter of ~930 nm in the (E) backward and (F) forward direction.

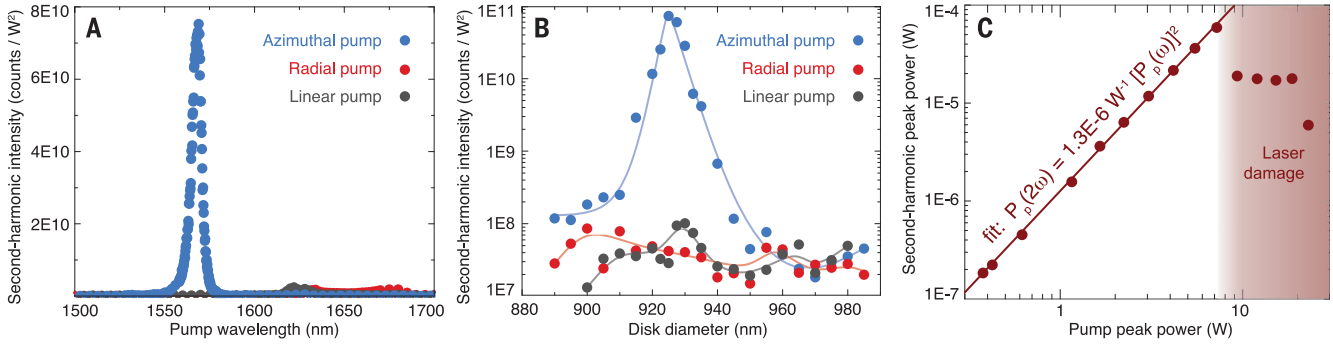


Fig. 4. Experimental nonlinear conversion efficiency. (**A** and **B**) Measured SH intensity as a function of the pump wavelength for the nanoresonator with the diameter of ~930 nm (A) and as a function of the nanoresonator diameter at a 1570-nm pump wavelength (B) for different pump polarizations. The SH intensity is normalized on the square of the pump power. (**C**) Measured peak SH power versus the peak pump power for a nanoresonator with a diameter of ~930 nm ($\log_{10}$ – $\log_{10}$ scale). Line shows the fit with a quadratic dependence with the nonlinear conversion coefficient 1.3 × 10^{-6} W^{-1}.

experimental parameters and an elaborated comparison with the earlier results for individual nanoresonators is presented in part 13 of the supplementary text (*25*). The SHG efficiency of an individual nanoantenna demonstrated here is qualitatively comparable to the best-to-date efficiencies of nonlinear metasurfaces (*29*, *30*) based on hybrid multiple-quantum-well structures, whereas a quantitative comparison cannot be done without some ambiguity. Although high nonlinear coefficients $P_p^{2\omega}/(P_p^{\omega})^2$ were demonstrated in such systems in the far- to mid-infrared spectral ranges, the reported conversion efficiencies $P_p^{2\omega}/P_p^{\omega}$ of 2 × 10^{-4}% (*29*) and 7.5 × 10^{-2}% (*30*), respectively, remain low and are limited by a peak pump power of 100 mW that they can sustain, compared with 10 W for our nanoresonator.

Our results illustrate, for the first time to our knowledge, manifestation of high Q-factor optical modes in individual nanoresonators in the linear and nonlinear regimes governed by the physics of bound states in the continuum.

Our experiments demonstrate that quasi-BIC engineering for individual nanoparticles in the optical frequency range is feasible despite fabrication tolerances and material absorption. Individual high-Q nanoresonators with a subwavelength footprint promise specific applications as nonlinear nanoantennas, low-threshold nanolasers, and compact quantum sources.

REFERENCES AND NOTES

1. A. I. Kuznetsov, A. E. Miroshnichenko, M. L. Brongersma, Y. S. Kivshar, B. Luk'yanchuk, *Science* **354**, aag2472 (2016).
2. A. Arbabi, Y. Horie, M. Bagheri, A. Faraon, *Nat. Nanotechnol.* **10**, 937–943 (2015).
3. D. Lin, P. Fan, E. Hasman, M. L. Brongersma, *Science* **345**, 298–302 (2014).
4. S. Kruk, Y. Kivshar, *ACS Photonics* **4**, 2638–2649 (2017).
5. S. S. Kruk *et al*., *Nano Lett.* **17**, 3914–3918 (2017).
6. V. F. Gili *et al*., *Opt. Express* **24**, 15965–15971 (2016).
7. J. Cambiasso *et al*., *Nano Lett.* **17**, 1219–1225 (2017).
8. J. Butet, P. F. Brevet, O. J. Martin, *ACS Nano* **9**, 10545–10562 (2015).
9. M. Celebrano *et al*., *Nat. Nanotechnol.* **10**, 412–417 (2015).
10. J. von Neumann, E. Wigner, *Phys. Z.* **30**, 465 (1929).
11. D. C. Marinica, A. G. Borisov, S. V. Shabanov, *Phys. Rev. Lett.* **100**, 183902 (2008).
12. C. W. Hsu, B. Zhen, A. D. Stone, J. D. Joannopoulos, M. Soljačić, *Nat. Rev. Mater.* **1**, 16048 (2016).
13. Z. F. Sadrieva *et al*., *ACS Photonics* **4**, 723–727 (2017).
14. Y. Plotnik *et al*., *Phys. Rev. Lett.* **107**, 183901 (2011).
15. C. W. Hsu *et al*., *Nature* **499**, 188–191 (2013).
16. K. Koshelev, Y. Kivshar, *Nature* **574**, 491–492 (2019).
17. A. Kodigala *et al*., *Nature* **541**, 196–199 (2017).
18. A. Tittl *et al*., *Science* **360**, 1105–1109 (2018).
19. F. Monticone, A. Alù, *Phys. Rev. Lett.* **112**, 213903 (2014).
20. M. G. Silveirinha, *Phys. Rev. A* **89**, 023813 (2014).
21. M. V. Rybin *et al*., *Phys. Rev. Lett.* **119**, 243901 (2017).
22. A. A. Bogdanov *et al*., *Adv. Photonics* **1**, 016001 (2019).
23. L. Carletti, K. Koshelev, C. De Angelis, Y. Kivshar, *Phys. Rev. Lett.* **121**, 033903 (2018).
24. M. P. Nezhad *et al*., *Nat. Photonics* **4**, 395–399 (2010).
25. See supplementary materials.
26. J. Wiersig, *Phys. Rev. Lett.* **97**, 253901 (2006).
27. J. van de Groep, A. Polman, *Opt. Express* **21**, 26285–26302 (2013).
28. J. A. Armstrong, N. Bloembergen, J. Ducuing, P. S. Pershan, *Phys. Rev.* **127**, 1918–1939 (1962).
29. J. Lee *et al*., *Nature* **511**, 65–69 (2014).
30. J. Lee *et al*., *Adv. Opt. Mater.* **4**, 664–670 (2016).

ACKNOWLEDGMENTS

We thank K. Ladutenko, B. Luther-Davies, D. Smirnova, and L. Wang for their valuable inputs into this project at various stages of its development. **Funding:** This work was supported by the Australian Research Council, the Strategic Fund of the Australian National University, the National Research Foundation of Korea (NRF) under grant no. 2018R1A3A3000666 funded by the Korean Government (MSIT), and the Russian Science Foundation under grant no. 18-72-10140. K.K. and A.B. acknowledge support from the Foundation for the Advancement of Theoretical Physics and Mathematics "BASIS." **Author contributions:** K.K., S.K., and Y.K. conceived the research; K.K. and A.B. performed theoretical analysis, numerical simulations, and data analysis; J.-H.C. and H.-G.P. fabricated the samples; S.K. and E.M.-G. conducted experimental studies; K.K., S.K., and Y.K. wrote the manuscript based on input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in this paper are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/288/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S15
Table S1
References (*31*–*43*)

5 September 2019; accepted 27 November 2019
10.1126/science.aaz3985

BIOMECHANICS

How flight feathers stick together to form a continuous morphing wing

Laura Y. Matloff[1], Eric Chang[1], Teresa J. Feo[2,3], Lindsie Jeffries[1], Amanda K. Stowers[1], Cole Thomson[1], David Lentink[1]*

Variable feather overlap enables birds to morph their wings, unlike aircraft. They accomplish this feat by means of elastic compliance of connective tissue, which passively redistributes the overlapping flight feathers when the skeleton moves to morph the wing planform. Distinctive microstructures form "directional Velcro," such that when adjacent feathers slide apart during extension, thousands of lobate cilia on the underlapping feathers lock probabilistically with hooked rami of overlapping feathers to prevent gaps. These structures unlock automatically during flexion. Using a feathered biohybrid aerial robot, we demonstrate how both passive mechanisms make morphing wings robust to turbulence. We found that the hooked microstructures fasten feathers across bird species except silent fliers, whose feathers also lack the associated Velcro-like noise. These findings could inspire innovative directional fasteners and morphing aircraft.

Bird flight feathers are hierarchically organized down to the micrometer scale (*1*–*6*), which makes them both firm enough to sustain lift and soft enough to smoothly overlap with adjacent feathers. Variable feather overlap enables birds to morph their wing and tail planforms more extensively than insects, bats (*7*), and current aircraft (*8*, *9*), providing unparalleled flight control (*10*, *11*); yet, how feather motion is coordinated during wing extension and flexion is not fully understood (*12*, *13*). Previous researchers hypothesized that flight feather coordination could be facilitated in several ways. A morphological study in pigeons (*13*) suggests that the smooth muscles and ligaments interconnecting the remiges may provide passive guidance for feather coordination. Feather microstructures termed "friction barbules" may prevent overlapping feathers from sliding too far apart during wing extension (*14*–*18*), but the mechanism responsible is unclear. Graham (*14*, *19*) suggested that the microscopic hooks of friction barbules may fasten adjacent feathers together by increasing friction, whereas subsequent work (*3*, *15*–*18*, *20*) suggested that they simply increase friction between feathers. Fastening and friction have different implications for our understanding of the evolution of avian flight. For instance, fastening during wing extension requires a mechanism to unfasten during flexion. On the other hand, "the energetic costs to overcome frictional forces during wing extension and flexion would be extremely large" (*12*).

To quantify how flight feathers are coordinated passively by means of elastic tissue between the base of the feathers, we measured the skeletal and flight feather kinematics of a rock pigeon (*Columba livia*) wing morphing between different glide poses (*21*, *22*) (Fig. 1, A and B; see methods). We found that feathers are redistributed through near-linear transfer functions that map the input wrist angle to each feather angle (Fig. 1, C and D). The slope represents the sensitivity of feather angle to wrist angle, and differences in slopes between adjacent feathers indicate how closely the motion of adjacent feathers is coupled (Fig. 1E). This shows how a series of tuned elastic ligaments between the remiges (the postpatagium) couples the wrist angle to all 20 remex angles (Fig. 1F). This 20:1 reduction in the number of degrees of freedom is classified as an underactuated mechanism in robotics, which formalizes earlier anatomical observations (*13*). However, it may not be entirely passive in vivo. Smooth muscles connecting the remiges may tune the stiffness of the underactuated mechanism (*13*), albeit not within a wingbeat cycle, because smooth muscles contract orders of magnitude more slowly (*23*, *24*). Although the corroborated elastic underactuation explains how feathers are distributed, it does not explain how gaps between feathers are prevented during wing extension.

When separating two overlapping pigeon flight feathers by hand, they initially slide smoothly before suddenly locking in place, suggesting that there must be a micromechanical end stop. To investigate this, we pressed the vane surfaces together with a predefined

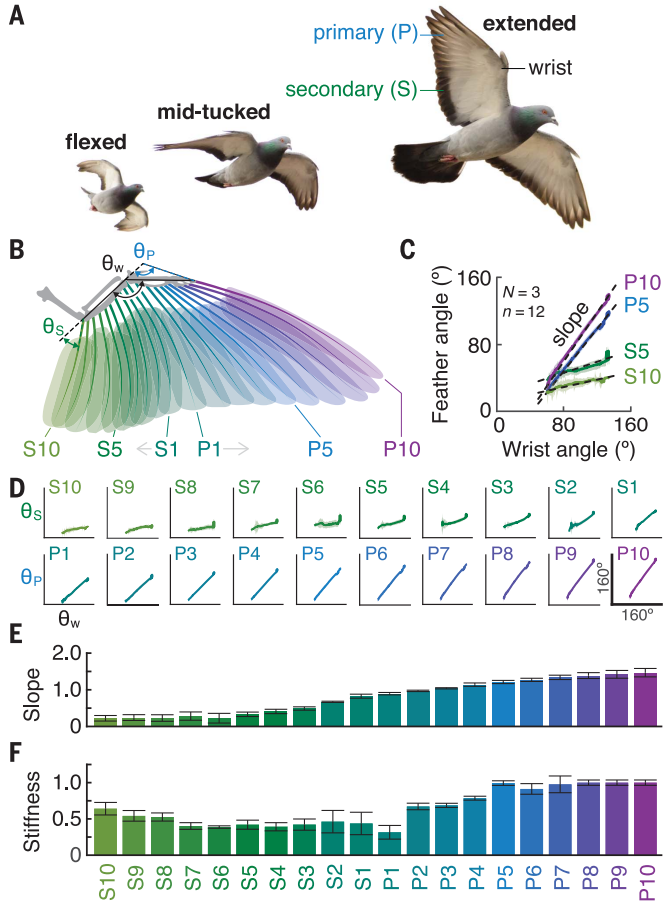


Fig. 1. Pigeon flight feathers are underactuated during wing flexion and extension. (**A**) Birds morph their wings during flight by flexing and extending their skeleton. (**B**) During morphing, as the wrist angle (θ_w) extends, flight feathers pivot relative to the ulna bone, measured by primary and secondary feather angles (θ_P and θ_S). (**C**) Linear transfer functions model the relationship between the wrist angle and feather angle. *N*, individuals; *n*, cycles each. (**D**) Measurements of all feather angles follow a linear relationship to wrist angle, suggesting underactuation. (**E**) The slopes of the linear model represent the sensitivity of the feather angles to wrist angle. (**F**) A linear elastic spring model corroborated from the feather transfer functions yields the normalized spring stiffness distribution of the connective tissue between the remiges. Error bars represent standard deviation.

[1]Department of Mechanical Engineering, Stanford University, Stanford, CA, USA. [2]Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA. [3]California Council on Science and Technology, Sacramento, CA, USA.
*Corresponding author. Email: dlentink@stanford.edu

normal force and then slowly rotated away the overlapping feather about the calamus tip using a computer-controlled motor. Simultaneously, we measured the time-resolved normal and opposing forces between the separating feathers (fig. S1 and methods). Across primary remiges (P10 and P9; P6 and P5), secondary remiges (S5 and S6), and rectrices (R5 and R6), we measured that flight feathers first slide with low opposing forces before they lock, causing the feathers to resist separation and the vanes to deform as a result (Fig. 2A). In the locked state, the force reaches a maximum, but because the feathers are forced to continue sliding, unfastening and refastening dynamically, they fail catastrophically (fig. S17 and movie S1) and separate (Fig. 2A). Simulating both wing flexion and extension, we observed that the opposing force is directional in pigeons: The maximum opposing force during extension is up to 10 times higher than during flexion (Fig. 2B). As a control, we slid the feather vanes along the rachis directions (anterior and posterior) and found low opposing forces similar to those in flexion. We evaluated the effect of normal force on the separation force as predicted by Coulomb's friction law: friction force = friction coefficient × normal force (Fig. 2C). A micrometer stage varied normal force by pressing feathers together from 50 mN

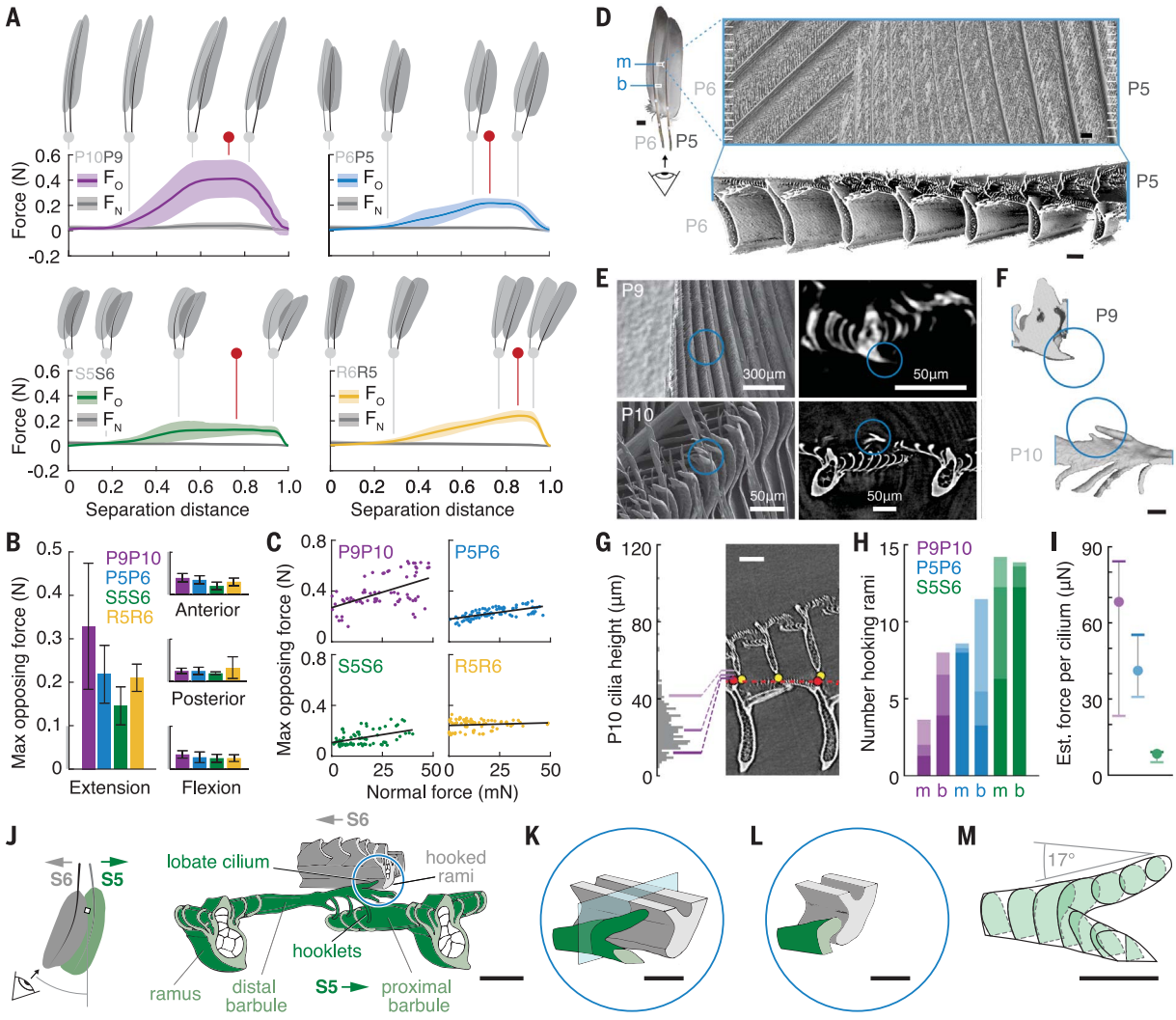


Fig. 2. Pigeon flight feathers lock together by means of microscopic directional probabilistic fasteners. (**A**) The opposing force (F_O) between pigeon flight feather pairs starts low and then ramps up before separation, whereas the normal force (F_N) remains low throughout (red pin indicates max force; values are averaged across normal force levels; feather outlines are from movie S1; shaded regions indicate force standard deviation). R, rectrix. (**B**) Maximal opposing forces are up to 10 times higher in the extension direction than in the flexion, posterior, and anterior directions. Error bars represent standard deviation. (**C**) Maximum opposing forces weakly depend on normal force and lack an intercept at zero. (**D**) Micro-CT scans of the overlapping feather pair P6-P5 show how their surfaces engage [scale bars are 10 mm (left) and 100 µm (top right and bottom right)]. See movie S2. m, middle; b, base. (**E** and **F**) Scanning electron microscopy images [(E), left], beamline micro-CT cross sections [(E), right], and three-dimensional reconstructions [(F); scale bar, 10 µm] of the microstructures (blue circles) involved in directional fastening. Top row: P9 overlapping outer vane rami with hook-shaped ventral ridge. Bottom row: P10 underlapping inner vane barbules with hooklike lobate dorsal cilia. (**G** and **H**) The distribution of lobate cilia protrusion height (G) was used to calculate the number of rami (yellow dots) hooked with cilia (beyond red dashed line; see methods for details) along vane-wise cross sections [white tick marks in (D)] (H). (**I**) Estimated force per hooked lobate cilium (see methods). Error bars represent standard deviation. (**J**) The interaction between a single lobate cilium and hooked ramus, as viewed from the feather tip (movie S3; scale bar, 50 µm). (**K** to **M**) The lobate cilium nestles snugly against the hooked ramus (L) via the sideward hooked lobe (M) after the slanted tip directs the ramus in position (scale bars, 10 µm; 17° is the angle for this lobate cilium).

down to as little as 0.2 mN (the sensor resolution limit). Even for the smallest normal forces, flight feathers locked in place with measured opposing forces close to 0.2 N (Fig. 2C and fig. S17). This large force, $^{1}/_{25}$ of body weight, is of the same order of magnitude as the lift that each flight feather has to support in gliding flight (body weight/40 remiges that form the wings). Coulomb friction requires unusually high friction coefficients, greater than 1000 for a normal force of 0.2 mN, well beyond established material properties (*25*). Furthermore, locking forces vary little with normal force and lack the intercept at zero normal force, which rules out Coulomb friction.

The distinctive interfeather fastener characteristics emerge from their hierarchical organization down to the microscale, which we visualized through scanning electron and x-ray microscopy (movies S2 and S3). Locking occurs in a spread wing or tail when the downward-curved outer vane of an overlapping flight feather slides across the upward-curved inner vane of an underlapping flight feather, in which the opposed curvatures ensure that the vane surfaces mate (*18*). In this region (Fig. 2D; see fig. S2 for nomenclature), the underlapping inner vanes have modified distal barbules with enlarged, hooked, lobate, dorsal cilia that extend above the dorsal ridge of the rami (Fig. 2E, bottom row; see figs. S3 to S6 for distributions). The overlapping outer vanes have rami with hook-shaped ventral ridges (Fig. 2E, top row, and figs. S9 and S10). To characterize the fastening mechanism between a hooked rami and single lobate cilium, we first estimated the number of locked lobate cilia in a feather pair (Fig. 2, G and H; fig. S12; and methods). The calculated maximum force per cilium is 10 to 70 μN (Fig. 2I), equivalent to the ~14 μN per hooklet (*4*) that zips barbs in the vane together (*3*, *4*, *6*). Notably, the same distal barbule functions both to fasten barbs within a flight feather by means of ventral hooklets and to fasten flight feathers within a wing by means of dorsal lobate cilia (Fig. 2J and fig. S13). The hooklets are oriented along the distal barbule to connect to proximal barbules, whereas the principal hooking direction of the lobate cilium is oriented to the side (Fig. 2, J to M, and fig. S11) to align with the hooked rami of the overlapping feather (figs. S7 and S8). Consequently, the principal hooking directions of the interbarb and interfeather fasteners are roughly orthogonal (Fig. 2J, fig. S13, and movie S3) and are thus functionally decoupled. The sophistication of the lobate cilium hooking mechanism culminates in its upward slanted tip sticking out above the rami (Fig. 2M and fig. S11). This enables the lobate cilium to catch and direct the overlapping hooked ramus so that its hooked lobe ends up nestling snug against the hooked ramus (Fig. 2, K and L, and movie S3), securely fastening both feathers during extension and automatically unlocking them during flexion. Fastening contradicts the hypothesized enhanced friction function of friction barbules (*3*, *12*, *14*–*20*), which we rename "fastening barbules" accordingly. This clarifies the function of the thousands of fastening barbules on the underlapping flight feathers; they lock probabilistically with the tens to hundreds of hooked rami of the overlapping flight feather and form a feather-separation end stop. The emergent properties of the interfeather fastener are not only probabilistic like bur fruit hooks, which inspired Velcro, but also highly directional like gecko feet setae (*26*)—a combination that has not been observed before (*27*).

To evaluate the function of both interfeather directional fastening and passive elastic feather redistribution on feather coordination in flight, we created a new biohybrid aerial robot with 40 underactuated pigeon remiges (Fig. 3A). We found that both underactuation and directional fastening are required to passively coordinate feather motion during dynamic wing morphing under calm outdoor flight as well as extremely turbulent conditions. Flight tests (Fig. 3B, fig. S14, and movie S4) demonstrated that the biohybrid wing morphs reliably at high flexion and extension frequencies of ~5 Hz, representative for pigeons (*21*). To quantitatively probe the function of passive elastic ligaments and interfeather directional fastening, we tested the robot wing at its approximate cruising speed (~10 m/s) and angle of attack (~10°) in a variable-turbulence wind tunnel. We manipulated the robot wing in four configurations in which we permutated removing the elastic ligaments and rotating the feathers along the rachis to separate the vanes (see methods). Tests in both high turbulence (30%; Fig. 3C) and low turbulence (3%; fig. S15) showed that elastic underactuation and feather fastening are required for continuous morphing. Without feather contact, but with elastic ligaments, gaps form between the primary feathers (Fig. 3D). Without elastic ligaments, but with feather contact, even larger gaps form as feathers move together in clumps (Fig. 3E). The wing without either elastic ligaments or feather contact has no coherent

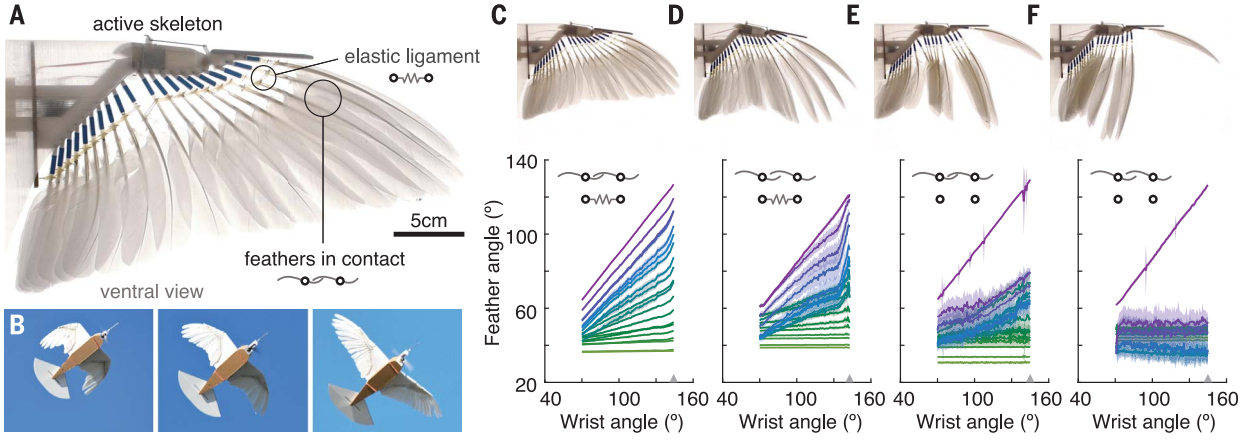


Fig. 3. Underactuated remiges with directional probabilistic fasteners morph robustly in flight. (**A**) We developed a biohybrid robotic wing with active skeletal control and 20 underactuated pigeon remiges in each wing half to evaluate the function of passive elastic ligaments and probabilistic directional fastening of adjacent feathers during dynamic wing morphing. (**B**) Successful outdoor flight of the biohybrid robot, demonstrating fully tucked, mid-tucked, and fully extended wings (movie S4). (**C** to **F**) Wind tunnel testing of the biohybrid wing in high turbulence (movie S5) with elastic coupling [(C) and (D)] and feather contact [(C) and (E)]. Both elastic coupling and feather contact (C) are required to maintain a continuous planform during wing morphing, whereas all other conditions [(D) to (F)] result in unnatural separation of the feathers and gaps in the planform. Note that the outermost feather P10 and the innermost feather S10 are fixed to the skeleton (for low turbulence, see fig. S15). Views are of the underside of the wing. Color scheme is the same as in Fig. 1B.

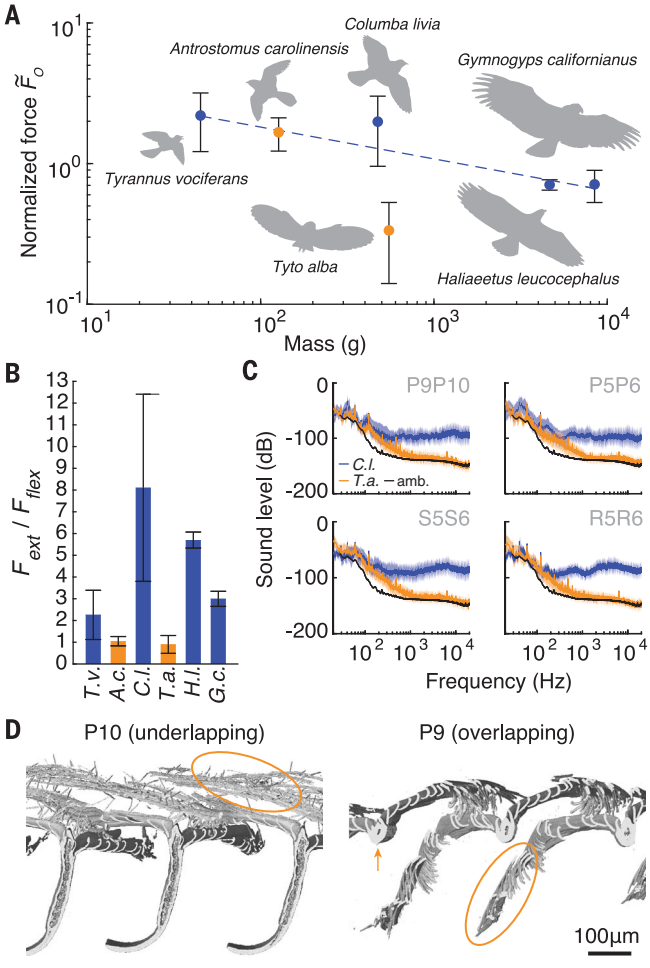


Fig. 4. Scaling and specialization of probabilistic feather fasteners across species. (**A**) Measured opposing feather-locking forces normalized by the nominal aerodynamic loading of each flight feather [$\tilde{F}_O = F_O$/(body weight/number of remiges)] for bird species ranging from ~40 g (Cassin's kingbird, *Tyrannus vociferans*) to ~9000 g (California condor, *Gymnogyps californianus*) (table S2). The trendline (blue dashed line) is shown with silent-flight species (orange) omitted. Silhouettes are based on fieldbook illustrations (*30*). Error bars represent standard deviation. (**B**) Extension-to-flexion ratios of opposing force (F_{ext} and F_{flex}) show that feather forces are directional, except for the species specialized in silent flight: barn owl (*T.a.*, *Tyto alba*) and chuck-will's-widow (*A.c.*, *Antrostomus carolinensis*). *T.v.*, *T. vociferans*; *C.l.*, *C. livia*; *H.l.*, *Haliaeetus leucocephalus*; *G.c.*, *G. californianus*. Error bars represent standard deviation. (**C**) Feather separation is much noisier in pigeon feathers than in owl feathers (movie S6). Shaded regions indicate standard deviation. amb., ambient noise level. (**D**) Beamline micro-CT scan of barn owl feathers shows the lack of lobate dorsal cilia in the inner vane of underlapping P10 and the lack of a hooked ventral ridge (orange arrow) in the outer vane rami of overlapping P9. Instead, both P10 distal barbules and P9 proximal barbules have elongated pennulae (~3-µm-diameter elongated structures in orange ellipses) that project beyond the plane of the rami.

feather coordination (Fig. 3F). In both outdoor flights and wind tunnel experiments, we consistently observed that directional fastening prevents gaps in the wing planform by locking adjacent remiges that risk separation, whereas elastic underactuation redistributes the locking forces to move unlocked remiges in place. Remiges only lock simultaneously during extreme wing extension, when feather sliding velocities approach zero (fig. S16), which minimizes the rate of energy loss (force × sliding velocity). Finally, the strong directionality in the locking force ensures that wing flexion is not resisted.

The directional probabilistic fastening mechanism between adjacent flight feathers is present across a wide range of modern bird species on the basis of three independent lines of evidence. First, the lobate dorsal cilia that enable fastening have been described across a wide range of species (*28*) (table S4). Second, we qualitatively observed interfeather fastening forces across a diverse set of species, except for silent flyers such as owls (table S3 and methods). Finally, we directly measured the interfeather fastening forces across select bird species ranging in body mass from a ~40-g Cassin's kingbird to a ~9000-g California condor (Fig. 4, A and B). The maximum measured force normalized by the estimated aerodynamic loading of each flight feather (body weight/number of remiges) has an order of magnitude of one across birds and scales only weakly with mass (mass$^{-0.2}$; Fig. 4A). Consequently, feather fastening forces are a similar fraction of body weight, and thus similarly effective, in both small and big birds. The fastening force is directional, with a force ratio of at least two between extension versus flexion across this range (Fig. 4B), except for the silent fliers (barn owl and chuck-will's-widow; Fig. 4B and table S3). High-resolution computerized tomography (CT) scans of barn owl feathers show that they indeed lack the lobate cilia and hooked rami in regions of feather overlap and instead have modified barbules with elongated, thin, velvety pennualue (Fig. 4D). This explains the low friction-like opposing forces we measured between their feathers (figs. S18A and S19A). Indeed, completely separating overlapping pairs of pigeon feathers produces a Velcro-like broadband sound, whereas separating barn owl flight feathers produces comparatively little noise, roughly 40 dB lower at 1 kHz (Fig. 4C, fig. S20, and methods). This confirms a functional trade-off between feather fastening and sound dampening (Fig. 4C), which Graham noted (*19*), and may explain the evolutionary loss of fastening barbules in species under selection for silent flight. We hypothesize that directional fastening may not be as critical for some silent fliers because decaying atmospheric turbulence at night (*29*) reduces the risk of feather slipping. The evolution of fastening barbules thus represents an important functional innovation in the transition from feathered dinosaurs to modern birds, which fossils may shed light on.

REFERENCES AND NOTES

1. R. O. Prum, *J. Exp. Zool.* **285**, 291–306 (1999).
2. M. Yu, P. Wu, R. B. Widelitz, C.-M. Chuong, *Nature* **420**, 308–312 (2002).
3. A. M. Lucas, P. R. Stettenheim, *Avian Anatomy: Integument Part I and Part II* (U.S. Government Printing Office, 1972).
4. A. Kovalev, A. E. Filippov, S. N. Gorb, *J. R. Soc. Interface* **11**, 20130988 (2013).
5. T. N. Sullivan *et al.*, *Adv. Funct. Mater.* **27**, 1702954 (2017).
6. F. Zhang, L. Jiang, S. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 10046–10051 (2018).
7. C. J. Pennycuick, *Modelling the Flying Bird* (Elsevier, 2008).
8. S. Barbarino, O. Bilgen, R. M. Ajaj, M. I. Friswell, D. J. Inman, *J. Intell. Mater. Syst. Struct.* **22**, 823–877 (2011).
9. D. D. Chin, L. Y. Matloff, A. K. Stowers, E. R. Tucci, D. Lentink, *J. R. Soc. Interface* **14**, 20170240 (2017).
10. D. Lentink *et al.*, *Nature* **446**, 1082–1085 (2007).
11. K. E. Crandell, B. W. Tobalske, *J. Exp. Biol.* **214**, 1867–1873 (2011).
12. J. J. Videler, *Avian Flight* (Oxford Univ. Press, 2006).
13. T. L. Hieronymus, *J. Anat.* **229**, 631–656 (2016).
14. R. R. Graham, *J. R. Aeronaut. Soc.* **36**, 24–58 (1932).
15. H. Sick, *J. Ornithol.* **85**, 206–372 (1937).
16. E. Rutschke, *J. Ornithol.* **106**, 307–312 (1965).
17. H. Oehme, *J. Ornithol.* **100**, 363–396 (1959).
18. H. Oehme, *Biol. Zent. Bl.* **82**, 569–587 (1963).
19. R. R. Graham, *Aeronaut. J.* **38**, 837–843 (1934).
20. N. S. Proctor, P. J. Lynch, *Manual of Ornithology: Avian Structure and Function* (Yale Univ. Press, 1993).
21. C. J. Pennycuick, *J. Exp. Biol.* **49**, 527–555 (1968).
22. A. K. Stowers, L. Y. Matloff, D. Lentink, *J. R. Soc. Interface* **14**, 20170224 (2017).
23. B. B. Hafen, B. Burns, in *Anatomy* (StatPearls Publishing, 2018); www.ncbi.nlm.nih.gov/books/NBK532857/.
24. J. C. Rüegg, *Physiol. Rev.* **51**, 201–248 (1971).

25. D. E. Gray, *American Institute of Physics Handbook* (McGraw-Hill Book Company Inc., 1957), section 2d.
26. K. Autumn *et al*., *Nature* **405**, 681–685 (2000).
27. S. N. Gorb, in *Springer Handbook of Nanotechnology*, B. Bhushan, Ed. (Springer, 2010), pp. 1525–1551.
28. A. C. Chandler, *A Study of the Structure of Feathers: With Reference to Their Taxonomic Significance* (Univ. California Press, 1916), vol. 13.
29. R. B. Stull, *An Introduction to Boundary Layer Meteorology* (Springer, 2012), vol. 13.
30. C. S. Robbins, B. Bruun, H. S. Zim, A. Singer, *Birds of North America: A Guide to Field Identification* (Golden Books Publishing Company Inc., 1983).
31. L. Y. Matloff, E. Chang, T. J. Feo, L. Jeffries, A. K. Stowers, C. Thomson, D. Lentink, How flight feathers stick together to form a continuous morphing wing online data. Figshare (2019); http://dx.doi.org/10.6084/m9.figshare.11369364.

ACKNOWLEDGMENTS

We thank A. Engilis and I. E. Engilis for feather loans; A. Vailiones for CT scanning help with NSF ECCS-1542152 support; K. Rogers (CDFW), K. Miller, and E. I. Knudsen for owl cadavers; J. Kadis, C. Basica, E. Georgieva, and B. Hightower for sound recording help; G. W. Reich for AFOSR flight permit support; and R. Prum for beamline 2-BM CT scanning of feathers with T.J.F. at Argonne National Laboratory (U.S. Department of Energy proposal 41887 to R. Prum) and analysis, with help of X. Xiao, S. Whittaker, C. Dove, and H. James for SEM support at the Smithsonian National Museum of Natural History. **Funding:** This project was supported by AFOSR BRI award number FA9550-16-1-0182 and AFOSR DESI award number FA9550-18-1-0525, with special thanks to B. L. Lee, F. A. Leve, and J. L. Cambier leading the program. T.J.F. was supported by NSF PRFB 1523857. E.C. and L.J. were supported by a NSF GRFP fellowship, A.K.S. by a NDSEG fellowship, and D.L. by NSF CAREER award 1552419. **Author contributions:** All authors contributed to the conception and design of the study and data analysis. D.L., L.Y.M., T.J.F., and E.C. wrote the manuscript; L.Y.M., E.C., T.J.F., L.J., and D.L. made the figures; L.Y.M. and A.K.S. performed motion capture; L.Y.M. made the spring model; L.Y.M. and C.T. performed feather interaction force measurements; T.J.F. performed scanning electron microscopy and beamline micro-CT imaging and analysis; L.J. performed micro-CT imaging of feather interaction; L.J., L.Y.M., and E.C. estimated force per cilium; E.C. fabricated the robot body; A.K.S., L.Y.M., and E.C. iterated and fabricated biohybrid robot wings; E.C. and L.Y.M. performed wind tunnel tests; E.C., L.Y.M., and A.K.S. performed flight tests; L.Y.M. and E.C. performed feather sound measurements; and D.L. contributed advice and supervised the work. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials, as well as online at figshare (*31*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/293/suppl/DC1
Materials and Methods
Figs. S1 to S20
Tables S1 to S4
References (*32*–*37*)
Movies S1 to S6

31 August 2019; accepted 18 December 2019
10.1126/science.aaz3358

MARTIAN ATMOSPHERE

Stormy water on Mars: The distribution and saturation of atmospheric water during the dusty season

Anna A. Fedorova[1]*†, Franck Montmessin[2]†, Oleg Korablev[1]†, Mikhail Luginin[1], Alexander Trokhimovskiy[1], Denis A. Belyaev[1], Nikolay I. Ignatiev[1], Franck Lefèvre[2], Juan Alday[3], Patrick G. J. Irwin[3], Kevin S. Olsen[2,3], Jean-Loup Bertaux[1,2], Ehouarn Millour[4], Anni Määttänen[2], Alexey Shakun[1], Alexey V. Grigoriev[1,5], Andrey Patrakeev[1], Svyatoslav Korsa[1], Nikita Kokonkov[1], Lucio Baggio[2], Francois Forget[4], Colin F. Wilson[3]

The loss of water from Mars to space is thought to result from the transport of water to the upper atmosphere, where it is dissociated to hydrogen and escapes the planet. Recent observations have suggested large, rapid seasonal intrusions of water into the upper atmosphere, boosting the hydrogen abundance. We use the Atmospheric Chemistry Suite on the ExoMars Trace Gas Orbiter to characterize the water distribution by altitude. Water profiles during the 2018–2019 southern spring and summer stormy seasons show that high-altitude water is preferentially supplied close to perihelion, and supersaturation occurs even when clouds are present. This implies that the potential for water to escape from Mars is higher than previously thought.

Mars once harbored an active hydrological cycle, as demonstrated by geological features on its surface, but it no longer holds the quantity of water required to produce such geological imprints (*1*, *2*). The planet's bulk inventory of water amounts to a global equivalent layer (GEL) of ~30 m, mostly contained in its polar ice caps (*2*). This is less than 10% of the water that once flowed on the surface (*1*). Mars' enhanced concentration of heavy water (semiheavy water five or more times the terrestrial standard) (*3*–*5*), strengthens the hypothesis that most of Mars' primordial water has escaped over time.

Water in the atmosphere is a negligible component of the planet's total water inventory, being equivalent to a global layer 10-μm thick, but nevertheless regulates the dissipation of water over time. Most martian water has been lost to space because its decomposition products (atomic hydrogen and oxygen) reach the upper atmosphere, where they can acquire sufficient thermal energy to overcome the low gravity of Mars (which is about one-third that of Earth's). Water decomposition is theorized to follow a complex reaction chain involving the recombination of H atoms into H_2 on a time scale of centuries (*6*–*8*), buffering any short-term hydrogen abundance variations. This mechanism has been challenged by observations showing that freshly produced H atoms can reach the exosphere (the uppermost layer where the atmosphere thins out and exchanges matter with interplanetary space) on a monthly time scale (*9*, *10*). The observed short-term variability of the hydrogen atoms populating the exosphere could be caused by direct deposition of water molecules at altitudes high enough to expose them to sunlight, which subsequently triggers a rapid enhancement of hydrogen atoms in the exosphere (*11*–*13*).

Testing this hypothesis requires characterizing the mechanisms contributing to upward water propagation through large-scale atmospheric circulation. One such mechanism is the cold trap imposed (as on Earth) by water ice cloud formation at low altitude, subsequent to water condensation. The condensation is predicted to occur whenever the partial pressure of water vapor exceeds saturation. The vapor pressure law causes the cold trap efficiency to depend heavily on temperature, which eventually limits the amount of water that can be transported to higher altitudes (*14*–*16*).

We investigate these processes using occultations of the Sun by the martian atmosphere (henceforth, solar occultations), where the vertical distributions of gases and particles can be directly observed. We used the Atmospheric Chemistry Suite (ACS) (*17*) on the ExoMars Trace Gas Orbiter (TGO) spacecraft. ACS is an assembly of three infrared spectrometers that together provide continuous spectral coverage from 0.7 to 17 μm, with a spectral resolving power ranging from 10,000 to 50,000. Our dataset was assembled by performing solar occultations with the near-infrared (NIR), mid-infrared (MIR), and thermal infrared in honor of professor V. I. Moroz (TIRVIM) channels of ACS. The NIR channel (0.7 to 1.7 μm) encompasses absorption bands of CO_2, H_2O, CO, and O_2, diagnostic of their molecular concentrations over altitudes of 5 to 100 km, with a vertical resolution of 1 to 3 km. TIRVIM (2 to 17 μm) provides simultaneous information on dust and water ice particle abundance.

We retrieved the volume fraction of water (i.e., its mixing ratio) and temperature using established methods (*18*–*20*), including the joint extraction of CO_2 and H_2O molecular abundances from the 1.57- and 1.38-μm absorption bands, respectively. Spectra were fitted with a spectroscopic model at all altitudes below 100 km, and the profiles of gaseous components were subsequently retrieved using an iterative algorithm (*21*). Figure 1A shows an example of model outputs fitted to the spectra, along with the resulting vertical water vapor profile. The sensitivity to H_2O depends on altitude, as it is a strong function of both the total number of molecules along the line of sight and the signal-to-noise ratio (SNR). SNR decreases exponentially with increasing atmospheric opacity along the line of sight, which is dominated by suspended dust aerosols and icy particles. Sensitivity to water vapor reaches 0.1 parts per million by volume (ppmv) at low altitudes in clear atmospheric conditions and is typically better than 1 ppmv between 10 and 75 km, rising up to ~20 ppmv at 100 km (*17*, *21*).

ACS NIR resolves the spectral structure of the CO_2 rotational band, providing simultaneous temperature and pressure parameters self-consistently (*21*). This simultaneity allows us to evaluate the local water vapor saturation state (Fig. 1B), a necessary parameter to estimate how much water can pass through the condensation level and reach the upper atmosphere. For most occultations, the NIR water vapor and temperature profiles can also be evaluated against aerosol profiles from ACS TIRVIM or ACS MIR data (Fig. 1B and fig. S1) (*21*).

The Sun-synchronous near-polar orbit of the TGO allows us to survey water vapor and aerosol vertical distributions on a global scale. The TGO performs two occultations on each 2-hour orbit. On average, ACS NIR accomplished nine occultation observations per sol (martian day) in both hemispheres [except during five periods of ~15 to 20 days each, around solar longitudes L_s 175°, 205°, 270°, 330°, and 355° (*22*)]. This produced a dataset of ~1700 occultations between April 2018 (L_s 163° MY34, where MY stands for martian year) and March 2019 (L_s 356° MY34). A summary of our results is shown in Fig. 2.

[1]Space Research Institute of the Russian Academy of Sciences (IKI RAS), Moscow, Russia. [2]Laboratoire Atmosphères Milieux Observations Spatiales (LATMOS), Université Paris-Saclay, Sorbonne Université, Centre National de la Recherche Scientifique, Guyancourt, France. [3]Physics Department, Oxford University, Oxford, UK. [4]Laboratoire de Météorologie Dynamique, Sorbonne Université, Centre National de la Recherche Scientifique, Jussieu, Paris, France. [5]Research School of Astronomy and Astrophysics and Advanced Instrumentation and Technology Centre at Mount Stromlo Observatory, Australian National University, Canberra, Australia.
*Corresponding author. Email: fedorova@iki.rssi.ru
†These authors contributed equally to this work.

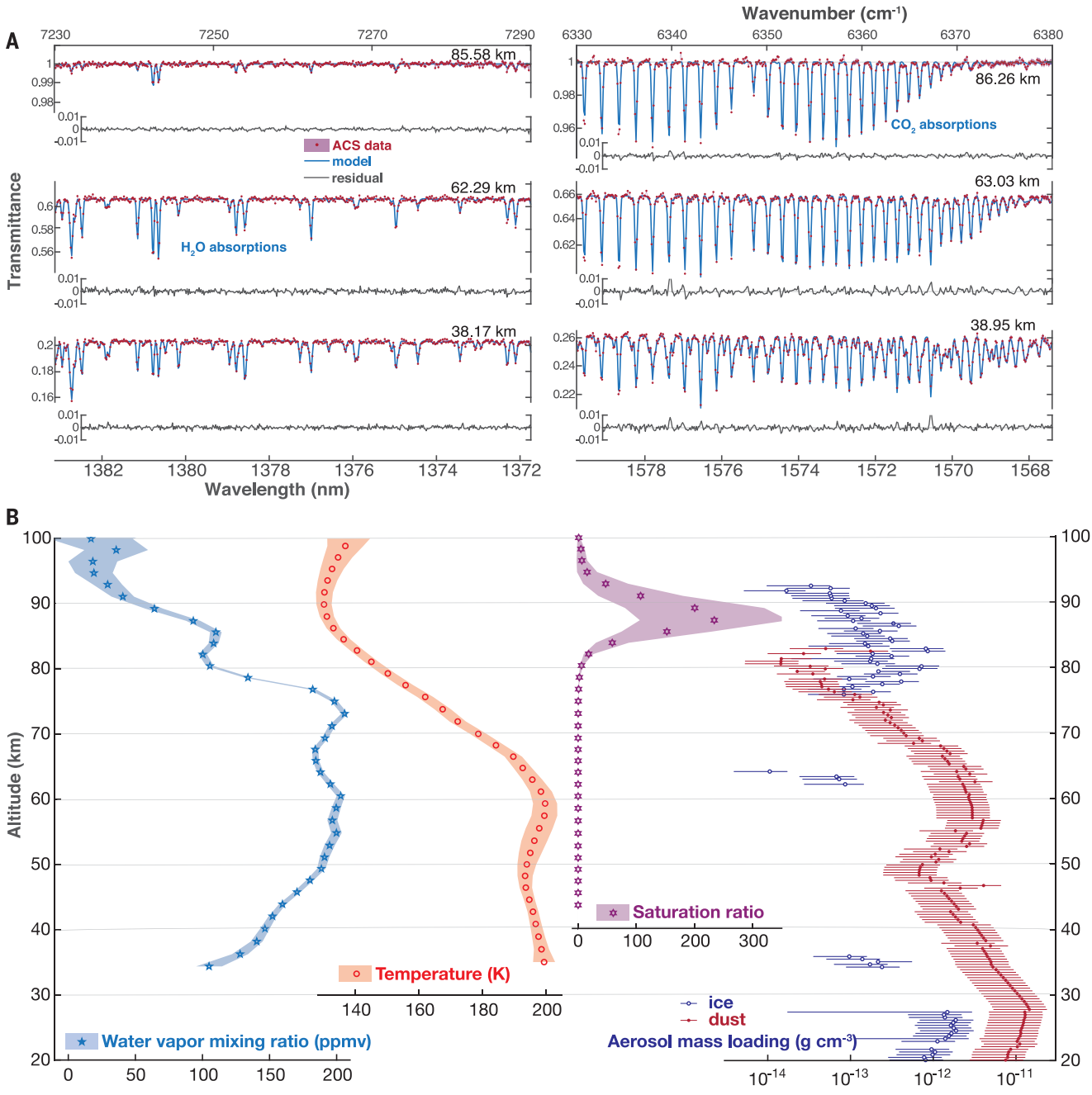


Fig. 1. Example ACS spectra and retrieved profiles. (**A**) ACS NIR spectra measured during orbit 2580 (L_s = 197.8°, 47.27°N, 85.2°W, local time 17:27) at three example altitudes, labeled in each panel. Synthetic models (blue curves) fitted to the data (red dots) account for the water content, CO_2 number density, and atmospheric temperature. The residuals are shown with gray lines. (**B**) Corresponding retrieved profiles of the H_2O mixing ratio, temperature, and saturation ratio derived from them. The profiles of the mass of aerosol particles (ice, blue circles; dust, red dots) per cubic centimeter obtained during the same orbit from ACS TIRVIM data are also shown. All altitudes are above areoid (equipotential surface for Mars, the analog of geoid for Earth). See (*21*) for details of the model fitting and retrieval procedures. The error bars or shaded regions, corresponding to 1σ confidence, are also explained in (*21*).

During MY34 southern spring and summer, the atmosphere was affected by two large-scale dust storms. The first, storm 2018A (L_s 188° to 250°) (*23*), enshrouded the planet globally [a global dust storm (GDS)], and the second, a recurrent large regional storm designated C (*24*, *25*) occurred later in the southern summer (L_s 320° to 335°).

Several prominent features are visible in Fig. 2. Both hemispheres exhibit relatively high water vapor mixing ratios (VMRs) throughout the southern spring-summer season. The southern hemisphere is distinctly wetter, with VMR exceeding 50 ppmv in the 50-to-100-km altitude range, while above 40 to 50 km, the northern hemisphere exhibits a gradual decline in VMR after L_s 210°. Water vapor propagates recurrently up to 100 km in the south, around L_s 200° (during the GDS), 270°, and 290°. In the north, this only occurred during the GDS (see also fig. S2). This migration of water vapor to high altitudes does not correlate with an increase in temperature, as evident around L_s 195° in the north, where water vapor propagates

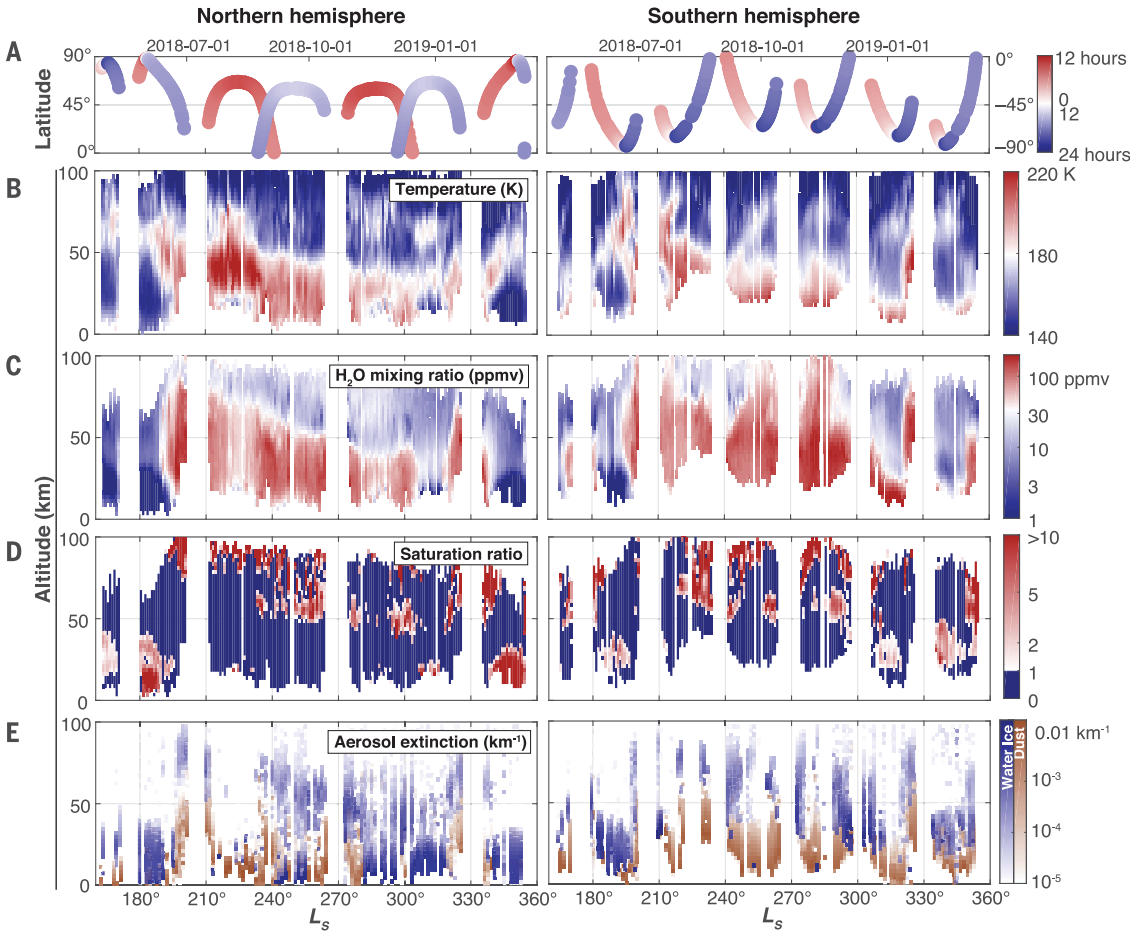


Fig. 2. Derived atmospheric properties during the dusty season of MY34. Each panel shows the data value in color, plotted as functions of L_s and altitude, with the northern hemisphere on the left and the southern hemisphere on the right. (**A**) Distribution of the ACS-NIR solar occultation observations, showing morning (red) and evening (blue) events. (**B**) Atmospheric temperature. (**C**) Water vapor mixing ratio. (**D**) Saturation ratio of water vapor; the blue regions correspond to an undersaturated state (i.e., saturation ratio < 1). (**E**) Water ice (blue) and dust (brown) aerosol extinctions.

higher than altitudes where the temperature rises. In the southern hemisphere, this is also evident around perihelion, when Mars orbits closest to the Sun. This implies that water vapor at high altitudes is primarily controlled by the large-scale motion of the martian atmosphere, particularly the upward branch of the cross-hemispheric atmospheric circulation cell, also known as the Hadley cell.

The saturation state in each hemisphere is shown in Fig. 2D, along with the temperature (Fig. 2B), water vapor (Fig. 2C), and aerosol concentrations (Fig. 2E). Large saturation ratios (1 to 10) are present in both hemispheres below 30 km, before and during the early phase of the GDS (L_s < 190°). ACS was observing the latitudes poleward of 60° at that time, implying that at least a third of the global atmosphere between 5 and 30 km was supersaturated. This is reminiscent of a previously observed supersaturation phenomenon (*19*) yet extends over a greater vertical range. Similar features can be seen in both hemispheres during this period—thick water ice clouds in a supersaturated atmosphere—suggesting that cloud formation does not limit the atmospheric saturation even when dust aerosols, on which ice crystals can form, are present. The clearing of dust particles carried by falling ice crystals (known as scavenging) is therefore not the sole reason for the existence of supersaturation on Mars, as was previously proposed (*19*).

During the following season, the northern hemisphere exhibits saturation ratios of >10 after L_s 330°, perhaps beginning around L_s 315°, before the early phase of the C storm. The southern hemisphere exhibits a more prevalent supersaturation at the same time, in the form of a discrete supersaturated area stretching between 15- and 40-km altitude, with unsaturated air beneath it extending to the pole. The same feature reforms soon after the C storm but slowly vanishes by the time of the spring equinox (L_s 360°).

In the 80-to-100-km part of the profiles shown in Fig. 2D, a supersaturated layer seems to persist throughout the observing period. Although the saturation state is generally less reliable in that altitude range (fig. S1), this layer is distinctly observed at L_s 190° to 200° in the northern hemisphere (Fig. 1B), and water ice clouds are also observed. The existence of such high-altitude supersaturation indicates efficient ascent of water to the upper atmosphere. Other regions of saturation are observed intermittently in both hemispheres at 50 to 60 km. The temperature is erratic in this region owing to atmospheric waves that generate higher temporal variability, causing fluctuations of >20 K over a few weeks. Because this region was previously filled with humid air during the GDS, cloud formation is enhanced. The clouds form in a similar configuration to the lower troposphere, as discussed above, characterized by

the simultaneous presence of clouds and supersaturation. An alternative version of Fig. 2, showing the averaged VMR and saturation ratio profiles binned into three narrow L_s intervals, is shown in fig. S3.

Our analysis constrains the mechanism controlling water propagation from the lower to the upper atmosphere. While dust controls upward propagation of water during the 2018 GDS and the 2019 C storm, water vapor can effectively and persistently reach the upper atmosphere around perihelion in the southern hemisphere. This coincides with the seasonal intensification of the Hadley circulation that reaches its peak around L_s 240° and is active until L_s 290° (fig. S2).

Large portions of the atmosphere are in a state of supersaturation, complementing previous observations (*19*, *26*). Unconstrained by saturation, the water vapor globally penetrates through the cloud level, regardless of the dust distribution, facilitating the loss of water to space. Because supersaturation is observed concomitantly with dust or ice particles, we conclude that condensation does not efficiently prevent water vapor from becoming supersaturated, even when seeds for condensation exist. We speculate that this may be due to rapid drops in temperature and/or rises in water concentration, which occur faster than condensation can keep up with.

Our results also show that water access to high altitude is affected by the seasonal changes around perihelion. Although planetary-scale dust storms appear in this period, those irregular events have a lesser impact than does seasonal change, which we suggest is the major atmospheric regulator for water. The seasonal recurrence and duration of the perihelion climate dominate the intermittent and short-lived effects of nonperihelion storms. Because perihelion coincides with the most intense period of the Hadley circulation, whose upwelling region is theoretically located in the southern tropical latitudes (*27*–*30*), and with the warmest period of the year, the perihelion season has likely governed the escape of water to space over geological time scales.

REFERENCES AND NOTES

1. M. H. Carr, *Nature* **326**, 30–35 (1987).
2. J. Lasue *et al*., *Space Sci. Rev.* **174**, 155–212 (2013).
3. T. Owen, J. P. Maillard, C. de Bergh, B. L. Lutz, *Science* **240**, 1767–1770 (1988).
4. S. Aoki *et al*., *Icarus* **260**, 7–22 (2015).
5. G. L. Villanueva *et al*., *Science* **348**, 218–221 (2015).
6. M. B. McElroy, *Science* **175**, 443–445 (1972).
7. T. D. Parkinson, D. M. Hunten, *J. Atmos. Sci.* **29**, 1380–1390 (1972).
8. V. A. Krasnopolsky, *J. Geophys. Res. Planets* **107**, 5128 (2002).
9. M. S. Chaffin *et al*., *Geophys. Res. Lett.* **41**, 314–320 (2014).
10. J. T. Clarke *et al*., *Geophys. Res. Lett.* **41**, 8013–8020 (2014).
11. M. S. Chaffin, J. Deighan, N. M. Schneider, A. I. F. Stewart, *Nat. Geosci.* **10**, 174–178 (2017).
12. N. G. Heavens *et al*., *Nat. Astron.* **2**, 126–132 (2018).
13. V. A. Krasnopolsky, *Icarus* **321**, 62–70 (2019).
14. D. V. Michelangeli, O. B. Toon, R. M. Haberle, J. B. Pollack, *Icarus* **102**, 261–285 (1993).
15. R. T. Clancy *et al*., *Icarus* **122**, 36–62 (1996).
16. T. Navarro *et al*., *J. Geophys. Res. Planets* **119**, 1479–1495 (2014).
17. O. Korablev *et al*., *Space Sci. Rev.* **214**, 7 (2018).
18. A. A. Fedorova *et al*., *Icarus* **200**, 96–117 (2009).
19. L. Maltagliati *et al*., *Science* **333**, 1868–1871 (2011).
20. A. Fedorova *et al*., *Icarus* **300**, 440–457 (2018).
21. Materials and methods are available as supplementary materials.
22. The solar longitude L_s is the Mars-Sun angle, measured from the northern hemisphere spring equinox, which is defined as L_s = 0°. The northern winter solstice is at L_s = 270°. Perihelion, where Mars orbits closest to the Sun, occurs at L_s = 251°.
23. S. D. Guzewich *et al*., *Geophys. Res. Lett.* **46**, 71–79 (2019).
24. D. M. Kass, A. Kleinböhl, D. J. McCleese, J. T. Schofield, M. D. Smith, *Geophys. Res. Lett.* **43**, 6111–6118 (2016).
25. P. Wolkenberg *et al*., *Icarus* **310**, 32–47 (2018).
26. R. Todd Clancy *et al*., *Icarus* **293**, 132–156 (2017).
27. F. Forget *et al*., *J. Geophys. Res. Planets* **104**, 24155–24175 (1999).
28. M. I. Richardson, R. J. Wilson, *Nature* **416**, 298–301 (2002).
29. F. Montmessin, F. Forget, P. Rannou, M. Cabane, R. M. Haberle, *J. Geophys. Res. Planets* **109**, E10004 (2004).
30. D. S. Shaposhnikov, A. S. Medvedev, A. V. Rodin, P. Hartogh, *Geophys. Res. Lett.* **46**, 4161–4169 (2019).

ACKNOWLEDGMENTS

ExoMars is a joint space mission of the European Space Agency (ESA) and Roscosmos. The ACS experiment is led by the Space Research Institute (IKI) in Moscow, assisted by LATMOS in France. We thank the numerous people responsible for designing, building, testing, launching, communicating with, and operating the spacecraft and science instruments of the TGO. **Funding:** The project was funded by Roscosmos and Centre National d'Etudes Spatiales (CNES). The science operations of ACS are funded by Roscosmos and ESA. Authors affiliated with IKI acknowledge funding from the Russian government under grant number 14.W03.31.0017 and contract number 0120.0 602993 (0028-2014-0004). Authors affiliated with the University of Oxford acknowledge funding from the U.K. Space Agency under grants ST/R001502/1 and ST/P001572/1. Authors affiliated with LATMOS acknowledge funding from CNES and Centre National de la Recherche Scientifique (CNRS). K.S.O. acknowledges the Natural Sciences and Engineering Research Council of Canada grant PDF-516895–2018. **Author contributions:** A.A.F., F.M., and O.K. conceived of the study, collected input from the other authors, and wrote the paper. The ACS observations and raw dataset were prepared by A.T., A.V.G., A.S., A.P., and N.K. A.A.F. calibrated the NIR ACS data and analyzed the profiles (assisted by A.T. and J.-L.B.). N.I.I. provided TIRVIM calibrated data. A.T., K.S.O., and L.B. provided MIR calibrated data. D.B. provided the ACS MIR aerosol extinction profiles. M.L. provided retrieval of aerosol properties using TIRVIM, NIR, and MIR datasets. J.A. and P.G.J.I. provided the MIR temperatures for validations. S.K. validated the MCS dataset. F.M., F.L., F.F., E.M., A.M., and C.F.W. provided expertise on the chemistry, circulation, and microphysics (assisted by J.-L.B.). All authors contributed to the preparation of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The ACS data are available from ESA's Planetary Science Archive at https://archives.esac.esa.int/psa/#!Table%20View/ACS=instrument; we used the level 2 data (*21*). The temperature, H_2O, and aerosol extinction profiles retrieved from the ACS measurements and analyzed in this article are available at http://exomars.cosmos.ru/ACS_Results_stormy_water_vREzUd4pxG/, including the list of orbits we used.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/297/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S3
References (*31*–*49*)

1 August 2019; accepted 18 December 2019
10.1126/science.aay9522

CELL BIOLOGY

Lipid-gated monovalent ion fluxes regulate endocytic traffic and support immune surveillance

Spencer A. Freeman[1]*†, Stefan Uderhardt[2,3]*, Amra Saric[4]*, Richard F. Collins[1], Catherine M. Buckley[1,5], Sivakami Mylvaganam[1], Parastoo Boroumand[1], Jonathan Plumb[1], Ronald N. Germain[2], Dejian Ren[6], Sergio Grinstein[1,7]†

Despite ongoing (macro)pinocytosis of extracellular fluid, the volume of the endocytic pathway remains unchanged. To investigate the underlying mechanism, we used high-resolution video imaging to analyze the fate of macropinosomes formed by macrophages in vitro and in situ. Na^+, the primary cationic osmolyte internalized, exited endocytic vacuoles via two-pore channels, accompanied by parallel efflux of Cl^- and osmotically coupled water. The resulting shrinkage caused crenation of the membrane, which fostered recruitment of curvature-sensing proteins. These proteins stabilized tubules and promoted their elongation, driving vacuolar remodeling, receptor recycling, and resolution of the organelles. Failure to resolve internalized fluid impairs the tissue surveillance activity of resident macrophages. Thus, osmotically driven increases in the surface-to-volume ratio of endomembranes promote traffic between compartments and help to ensure tissue homeostasis.

During endocytosis, cells internalize membrane along with extracellular fluid (pinocytosis). The amount of fluid ingested can be substantial: dendritic cells and macrophages take up the equivalent of their entire cellular volume every 4 hours (*1*). Despite continuous uptake of large amounts of water and osmolytes, the endocytic pathway and the cells as a whole retain their volume and ionic composition over extended periods (*1*). To investigate endomembrane volume and ionic regulation, we chose macropinosomes, large (up to 5 μm) vacuoles formed by specialized cell types (*2*, *3*) (Fig. 1 and fig. S1, A and B). Unlike smaller endocytic vesicles, macropinosomes can be resolved by diffraction-limited microscopy (*4*), enabling detailed assessment of their volume as they mature. Moreover, the volume of medium entrapped by macropinosomes is sufficiently large to alter the overall ionic composition of the cells (fig. S1, C and D). When stimulated by macrophage colony-stimulating factor (M-CSF), bone marrow–derived macrophages (BMDM) underwent a large burst of macropinocytosis. Multiple large vacuoles (10 to 15 per cell; mean volume: 7 μm^3) formed within 5 min (Fig. 1A and fig. S1, A and B). The volume of fluid internalized was equivalent to ≈25% of the cell volume, an increase detectable by electronic cell sizing (Fig. 1B). When visualized using rhodamine-dextran (70 kDa), the macropinosomes of BMDM (Fig. 1A), and those formed by peritoneal macrophages and human monocyte–derived macrophages (fig. S1E), resorbed within 30 min, and cell volume returned to basal levels (Fig. 1B). Shrinkage of macropinocytic vacuoles was also observed in vivo. Two-photon imaging of live mice revealed that resident tissue macrophages (RTM) of the peritoneal serosa—interstitial, non-migratory cells that constitutively sample the surrounding milieu (*5*) (movie S1)—formed large macropinosomes that subsequently contracted (Fig. 1, C and D, and movie S2).

Macropinosome shrinkage was accompanied by an increase in the intensity of the luminal dextran fluorescence (Fig. 1A), implying that fluid was extracted from the vacuoles. This suggested that the volume loss of the vacuoles is caused by osmotically driven solvent loss. Na^+ and Cl^- constitute the majority of the osmolytes in the fluid engulfed during macropinocytosis. Accordingly, inducing macropinocytosis with M-CSF resulted in a fourfold increase in the total cellular Na^+ concentration (fig. S1D). Thus, loss of Na^+ and Cl^- along with osmotically coupled water may underlie the rapid shrinkage of macropinosomes. This was validated by ion substitution experiments: replacing Na^+ for the impermeant cation *N*-methyl-D-glucamine$^+$ (NMG^+) virtually eliminated macropinosome resolution (Fig. 1E; fig. S1, E and F; and movie S3). Similarly, shrinkage was precluded when substituting the impermeant anion gluconate$^-$ for Cl^-, implying that electroneutrality needs to be maintained during solute export (Fig. 1E and fig. S1, E and F). Preventing monovalent ion efflux from macropinosomes also prevented restoration of the cell volume (fig. S1G). The absence of luminal Ca^{2+} did not prevent macropinosome resolution (Fig. 1E).

We tested a series of ion transport inhibitors to gain insight into the pathways involved in macropinosome shrinkage. Tetrandrine, a potent inhibitor of two-pore channels (TPC) (*6*), impaired volume loss (Fig. 2A; fig. S2, A and B; and movie S4). Interestingly, the endomembrane isoforms TPC1 and TPC2 are expressed at particularly high levels in myeloid cells, including BMDM (fig. S6I) (*7*). Moreover, TPC1 is highly expressed in the macropinocytic interstitial ($Ccr2^-$, $CD169^+$) RTM, compared with neighboring stromal or migratory ($Ccr2^+$, $CD169^-$) myeloid cells that are nonmacropinocytic (Fig. 2, D and E) (*5*). Although undetectable at the plasma membrane, TPC1 was rapidly (in ≤1 min) acquired by nascent macropinosomes (Fig. 2F), whereas TPC2 was recruited later (fig. S3D). BMDM from *Tpc1;Tpc2* double-knockout mice formed large macropinosomes when stimulated by M-CSF, but these did not shrink and resolve during our analyses (Fig. 2, B and C). Using single knockout mice and RNA interference, we discerned this effect to be attributable primarily to TPC1 (Fig. 2C and fig. S3, F and G).

Certain ion channels, including TPCs and TRPMLs (mucolipin transient receptor potential channels), require phosphatidylinositol 3,5-bisphosphate [PtdIns(3,5)P_2] for activation (*8*, *9*). This phosphoinositide is generated by phosphorylation of PtdIns(3)P by the phosphoinositide kinase PIKfyve (*10*). PtdIns(3)P and PIKfyve itself were readily detectable on the cytosolic leaflet of nascent macropinosomes (Fig. 2F and fig. S3E), which is consistent with this sequence. 2xMLN–green fluorescent protein (GFP), a putative probe for PtdIns(3,5)P_2 (*11*) was also found in macropinosomes (fig. S3E). Macropinosome shrinkage—whether measured directly or assessed indirectly from the overall cell volume gain—was blocked by PIKfyve antagonists (Fig. 2, G to I, and fig. S1, L to N). Inhibiting PIKfyve did not alter the water permeability or pliability of the membrane, as indicated by the acute volume loss induced by water abstraction caused by hypertonic medium (fig. S1, K and L). A similar response to hypertonicity was observed in macropinosomes formed in Na^+-free solution (fig. S1J). Although TRPML1 channels are recruited to maturing macropinosomes, deletion of the *Trpml1* gene did not affect macropinosome resorption (Fig. 2C and fig. S3D).

The area of the vacuolar membrane was reduced during shrinkage by emission and severing of tubules and vesicles, which were visualized using FM 1-43 (Fig. 3A) or sulforhodamine B (Fig. 3D) dyes. Tubule extension accompanies and requires the volume loss that is driven by the export of ions and osmotically coupled water. Accordingly, substitution of Na^+ with NMG^+, or blockade of TPC channels by tetrandrine

[1]Program in Cell Biology, Peter Gilgan Centre for Research and Learning, Hospital for Sick Children, Toronto, ON, Canada. [2]Lymphocyte Biology Section, Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [3]Department of Internal Medicine 3 – Rheumatology and Immunology, Universitätsklinikum Erlangen and Friedrich-Alexander University Erlangen-Nürnberg (FAU), Erlangen, Germany. [4]Neurosciences and Cellular and Structural Biology Division, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA. [5]Institute of Microbiology and Infection and School of Biosciences, University of Birmingham, Edgbaston, Birmingham, UK. [6]Department of Biology, University of Pennsylvania, Philadelphia, PA, USA. [7]Keenan Research Centre of the Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, ON, Canada.
*These authors contributed equally to this work.
†Corresponding author. Email: spencer.freeman@sickkids.ca (S.A.F.); sergio.grinstein@sickkids.ca (S.G.)

or their inactivation by depleting PtdIns(3,5)P_2, impaired tubulation and vesiculation (Fig. 3, A and B). Genetic deletion of TPC channels also precluded tubulation (Fig. 3C). Applying a hypertonic solution to macropinosomes that initially failed to shrink because they were loaded with NMG^+ or formed in cells treated with a PIKfyve inhibitor revealed that tubulation is a consequence, not a cause, of volume loss (Fig. 3, D and E, and movie S5). The tubules emanating from early macropinosomes are very thin, with a modal diameter of ≈30 nm (Fig. 3F and fig. S4C), which likely explains the preferential retention and progressive concentration of large (70-kDa) dextran in the vacuolar lumen. The diameter of the tubules generated during macropinosome resorption makes them well suited for associating with proteins containing BAR domains, concave structures that preferentially bind and stabilize curved membranes of ≈22-nm diameter (*12*). Indeed, the BAR domain–containing proteins SNX1, SNX2, and SNX5 decorated the tubules emanating from resorbing macropinosomes (fig. S4A) (*13*). Thus, membrane crenation caused by the volume loss potentially generated the necessary curvature to stabilize BAR domains on the membrane and thereby fostered tubulation (Fig. 3I). This notion was tested by generating liposomes and monitoring the effects of hydrostatic tension on the ability of a recombinant BAR-domain protein, BIN1, to induce tubulation. These experiments revealed a distinctive relationship between volume loss and BAR-mediated tubulation: the relief of hydrostatic tension greatly amplified tubulation by BIN1, whereas swelling the liposomes counteracted it (Fig. 3G and fig. S4D). Given their functional redundancy and ability to form interchangeable heteromeric complexes, the loss of any one BAR protein is unlikely to prevent tubulation. Because many BAR-domain proteins (including various SNX isoforms) require PtdIns(3)P for optimal binding, we interfered with their association by scavenging the available head groups of the phosphoinositide by expressing tandem FYVE domains. Indeed, high-affinity (multicopy) FYVE-domain tandems prevented SNX recruitment and precluded tubulation and resolution of the macropinosomes (Fig. 3H and fig. S4B).

The significance of endomembrane shrinkage driven by efflux of monovalent ions is far-reaching. The resulting tubulation mediates the recycling of plasmalemmal components that are internalized in the course of macropinocytosis and endocytosis. One such example is Mac-1 (αMβ2), an integrin that is key to macrophage adherence, migration, and phagocytosis (*14*). Blocking TPC channels by inhibiting PtdIns(3,5)P_2 formation led to a pronounced depletion of plasmalemmal Mac-1, which was instead trapped in endomembrane vacuoles (Fig. 4A and fig. S5C). The effect was phenocopied by simply substituting extracellular Na^+ by K^+, demonstrating that a Na^+ gradient is

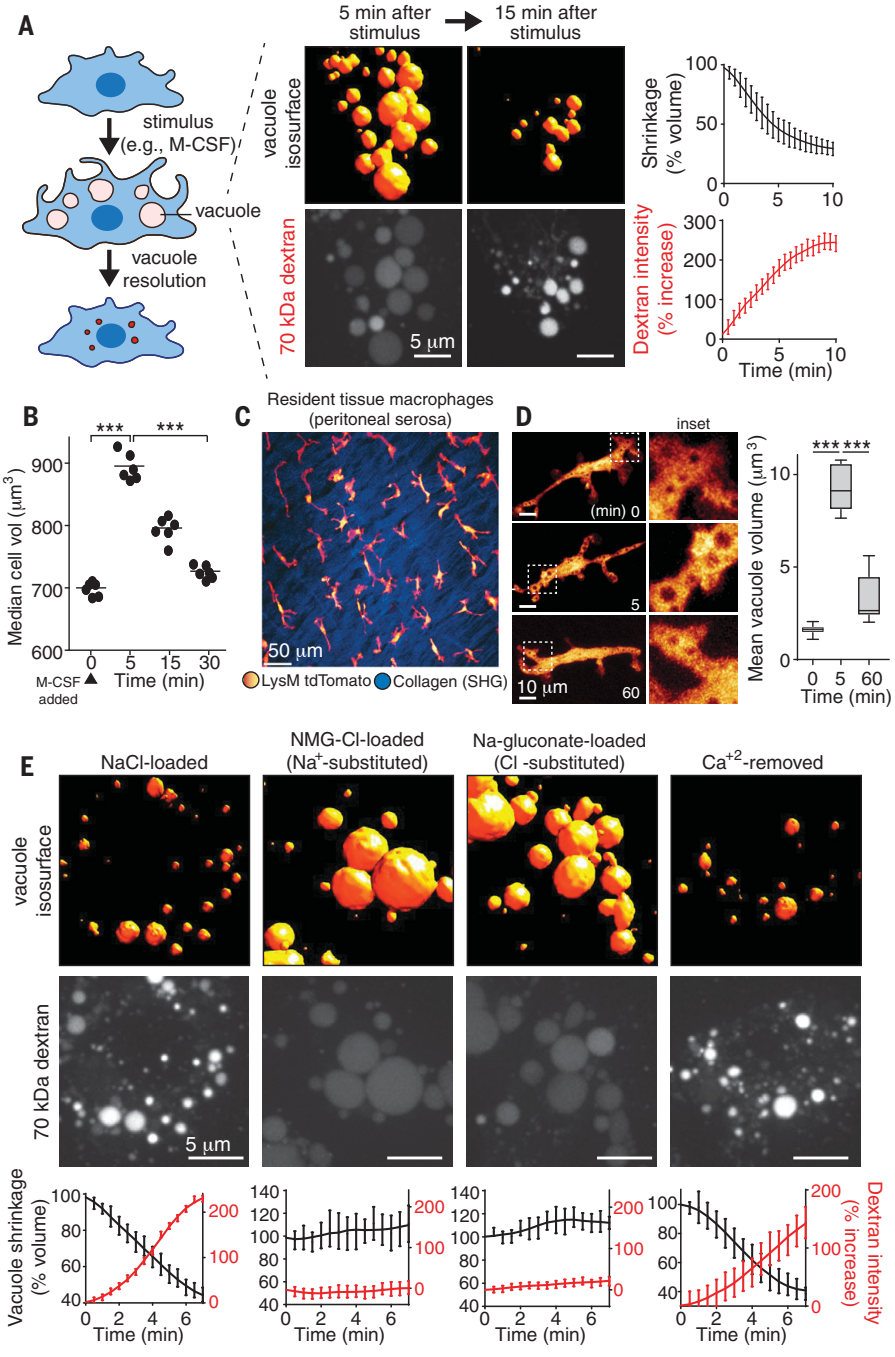


Fig. 1. Vacuolar shrinkage requires monovalent ion efflux. (**A**) Volume (vol) and 70 kDa rhodamine-dextran fluorescence intensity changes of macropinosomes induced in bone marrow–derived macrophages (BMDM) by macrophage colony-stimulating factor (M-CSF); data are means ± SEM of >100 vacuoles from three independent experiments (i.e., n = 3). Measurement of vacuole resolution was initiated after a 5-min stimulation with M-CSF in medium containing dextran, followed by an immediate wash. (**B**) Cell (BMDM) volume was measured electronically before and at the indicated times after M-CSF stimulation; $>10^4$ cells per point, n = 3. P values determined by unpaired, two-sided t tests. Here and elsewhere, ***P < 0.001, **P < 0.01, and *P < 0.05. (**C**) Intravital observation of td-Tomato–labeled resident tissue macrophages (RTM; pseudocolored yellow and red) of the peritoneal serosa and second harmonic imaging (SHG) of collagen (blue). See also movies S1 and S2. (**D**) Visualization and volume quantification of M-CSF–induced macropinosomes in RTM in vivo; means, upper and lower quartiles (boxes), and distribution (whiskers) are graphed. >50 vacuoles, n = 3. P values determined by Mann-Whitney U test. (**E**) Macropinosomes of M-CSF–stimulated BMDM in media containing indicated solutes and dextran. Representative images acquired at 5 min. See also movie S3. Bottom row: mean ± SEM macropinosomal volume and dextran intensity from three independent video recordings representing >150 macropinosomes. See also fig. S1.

Fig. 2. Monovalent ion efflux mechanisms. (**A**) Macropinosome volume changes in presence of 5 μM tetrandrine, measured in BMDM. Measurement of vacuole resolution was initiated once cells were washed after a 5-min stimulation with M-CSF in medium containing dextran; tetrandrine or vehicle were present throughout. Means ± SEM, *n* = 3. See also fig. S2 and movie S4. (**B** and **C**) Macropinosome volume changes after stimulation with M-CSF in WT, *Tpc1* and *Tpc2* single and double knockout (KO), and *Trpml1* KO BMDM. In (C), means, upper and lower quartiles (boxes), distribution (whiskers), and observations from fields containing three to five cells (dots) each, measured 10 min after macropinosome formation; *n* = 3. (**D**) Staining of the peritoneal serosa. Outline of CD169 signal (left) overlaid on TPC1 signal (right). (**E**) RNA sequencing. Resident tissue macrophages were Cx3cr1^{-}/Ccr2^{-}. Migratory cells were Cx3cr1^{-}/Ccr2^{+}. (**F**) BMDM expressing TPC1-tomato or 2xfyve-GFP to detect PtdIns(3)P. Dextran shown in cyan. (**G**) BMDM stimulated with M-CSF in the presence of 70 kDa rhodamine-dextran and, where indicated, the PIKfyve inhibitors YM201636, apilimod, or WX8 (all used at 500 nM). Resolution was recorded as in (A). 5 min after isosmotic recording, a hyperosmotic solution [final 500 milliosmolar (mOsm)] was added to verify the osmotic responsiveness of the vacuoles. (**H** and **I**) Visualization and volume quantification of macropinosomes in RTM treated in situ with YM201636 (500 nM) or tetrandrine (5 μM); >15 cells, *n* = 3. All *P* values determined by Mann-Whitney *U* tests.

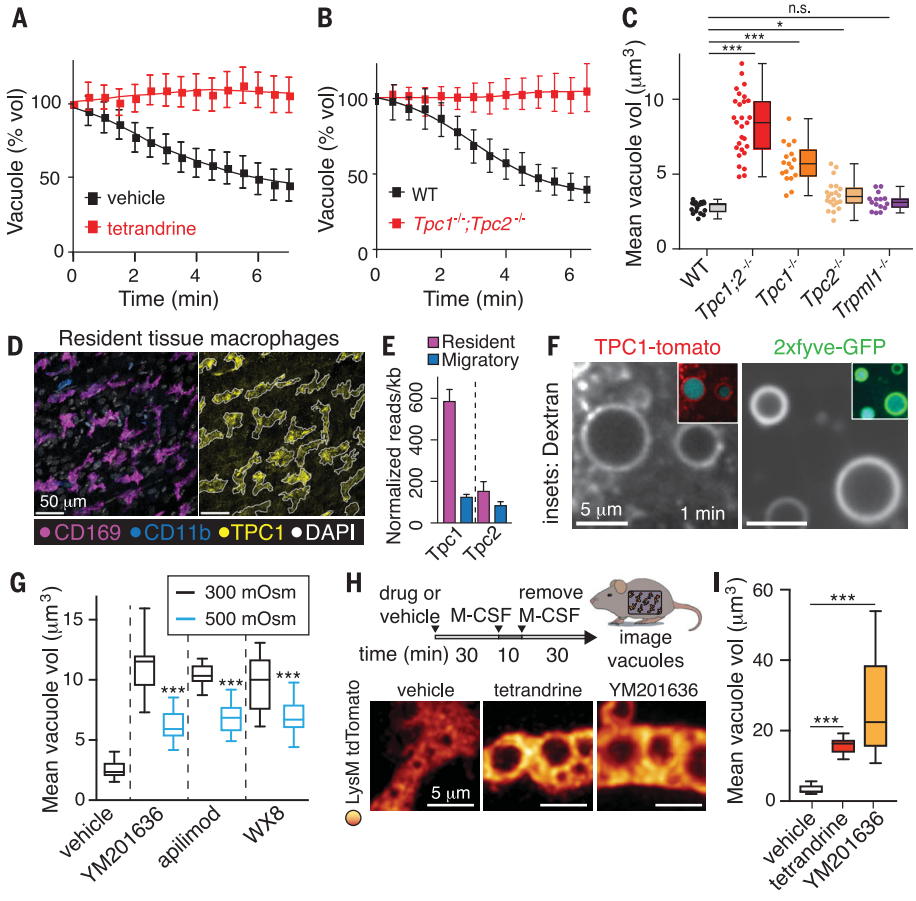
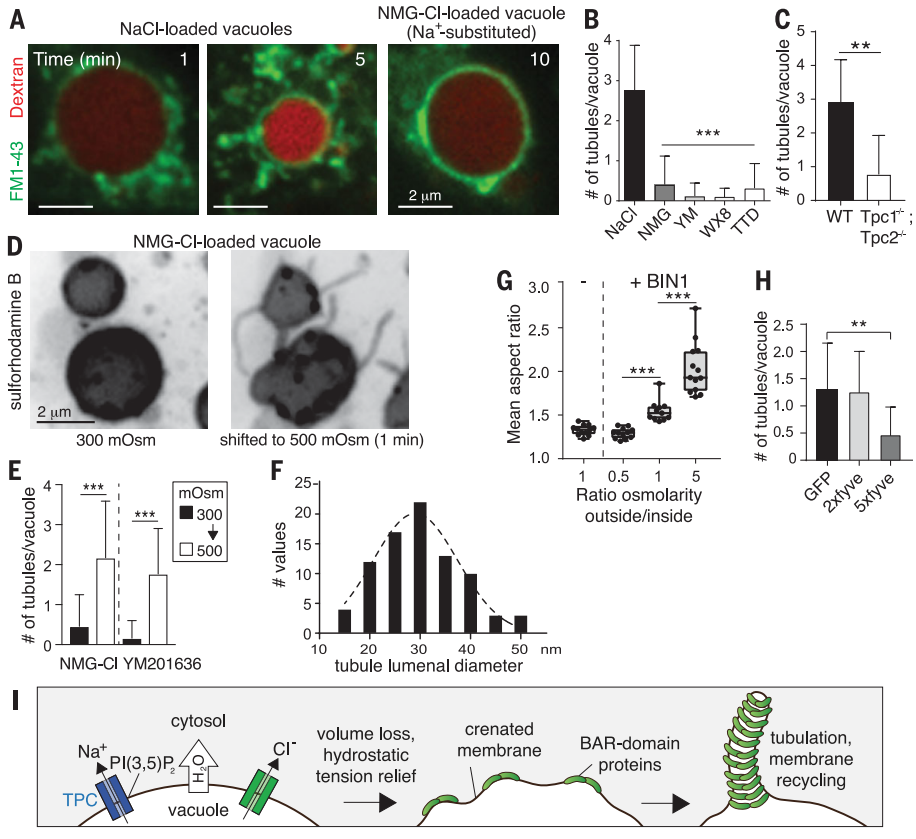


Fig. 3. Osmotically driven shrinkage induces tubulation. (**A**) BMDM were stimulated with M-CSF and the distribution of 70 kDa rhodamine-dextran and FM 1-43 imaged at the indicated times after removal of the stimulus. (**B** and **C**) Mean number of tubules (exceeding 1 μm in length) measured 5 min after stimulation with M-CSF; >100 macropinosomes (*n* = 3) for each condition. (**D** and **E**) Macropinosomes containing sulforhodamine B (SRB) and *N*-methylglucamine chloride or formed in cells treated with YM201636 (500 nM) were recorded 10 min after formation, before and after being subjected to hypertonic solution. See also movie S5. >100 vacuoles, *n* = 6. (**F**) Transmission electron microscopy was used to measure the diameter of tubules emerging from macropinosomes; 85 tubules were quantified. (**G**) Liposomes formed of whole brain lipid, rhodamine-labeled phosphatidylethanolamine, and PtdIns(4,5)P_2 in 20 mOsm solution. The mean aspect ratio for three to five fields of liposomes incubated with or without recombinant human BIN1 was quantified by imaging, *n* = 3. (**H**) HT1080 cells expressing GFP, 2xfyve-GFP, or 5xfyve-GFP were pulsed with SRB for 10 min. Mean number of tubules (exceeding 1 μm in length); >100 macropinosomes (*n* = 3) for each condition. All *P* values determined by unpaired, two-sided *t* tests. (**I**) Model of mechanism proposed to underlie macropinosomal tubulation.



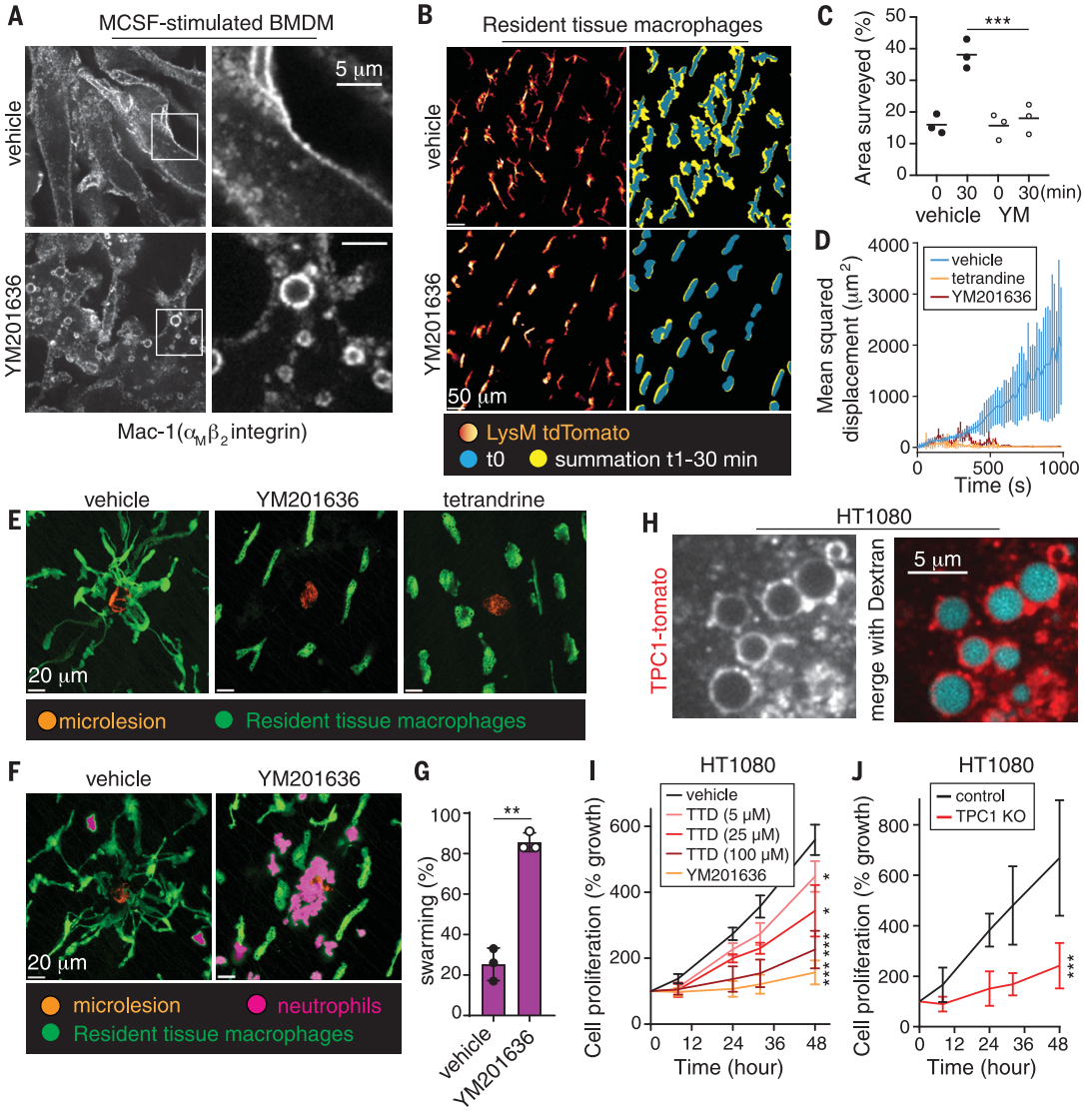


Fig. 4. Vacuolar resolution maintains cellular responsiveness and tissue surveillance. (**A**) BMDM were stimulated for 30 min with M-CSF, with or without YM201636, fixed and immunostained for Mac-1. (**B** to **E**) Resident tissue macrophages (LysM-tdTomato) with or without YM201636 or tetrandrine were stimulated with M-CSF for 10 min followed by removal of the stimulus for 30 min. Cells were then imaged for 30 min. (C) surveillance area measured in the absence (left) or presence (right) of YM201636 (YM) over time, $n = 3$. In (D and E), 30 min after M-CSF removal, laser-induced microlesions were generated, marked by the resultant autofluorescence [orange in (E)]. Mean squared displacement of the macrophages is graphed in (D); means ± SD, $n = 3$. Representative images in (E) taken at 15 min after injury. (**F** and **G**) Experiments performed as in (E). Representative images denoting the presence of neutrophils are shown. In (G), the percentage of lesions with neutrophil swarming was quantified for six microlesions per animal, $n = 3$. (**H**) HT1080 cells expressing TPC1-tomato were incubated with 70 kDa dextran for 10 min before imaging. (**I** and **J**) WT and TPC1 KO HT1080 cell growth measured in the absence or presence of YM201636 or tetrandrine (TTD) by cell counting. Means ± SD, $n = 3$. All P values determined by Mann-Whitney U tests.

exploited by the endocytic pathway to direct membrane traffic (fig. S5, A to C). Moreover, nonmacropinocytic cells (e.g., fibroblasts) also required Na^+ efflux to execute canonical receptor recycling (fig. S6, A to D, G, and H) (*15*, *16*). Defective plasmalemmal protein recycling caused by ion substitution had acute functional consequences: the ability of BMDM to bind and ingest complement-coated targets and of fibroblasts to form focal adhesions—processes mediated by integrins—were severely depressed (figs. S5, D and E, and S6, E and F).

The need for ion efflux from endocytic compartments for normal cell function was also documented in situ. The ability of interstitial RTM to survey their environment was impaired when Na^+ efflux from the endocytic pathway was prevented in vivo (Fig. 4, B to E, and movie S6). Blocking β1 and β2 integrins similarly inhibited surveillance (fig. S7). When microlesions are made by targeted laser ablation to adjacent fibroblasts, RTM normally emit processes to contain the damage (*5*), preventing the recruitment and activation of neutrophils. When PIKfyve or TPCs were inhibited, the cells failed to resorb vacuoles and were unable to respond to the damage (Fig. 4, D and E). As a result, neutrophil swarming ensued (Fig. 4, F and G). Thus, vacuole resolution, mediated by lipid-gated Na^+ efflux, underpins membrane traffic necessary to maintain cellular responsiveness.

Because macropinocytosis is not restricted to phagocytes, inhibition of vacuolar shrinkage also affects other cell types (*17*). HT1080 fibrosarcoma cells display vigorous constitutive macropinocytosis (Fig. 4H), express TPC1 at comparatively high levels (fig. S8A), and localize the channel to macropinosomes (Fig. 4H). HT1080 cells require growth factors for survival and are highly responsive to epidermal growth factor (EGF). Because the EGF receptor is internalized along with the macropinosomal membrane, effective recycling to the cell surface is key. Preventing vacuole shrinkage resulted in the depletion of EGF receptors from the plasmalemma (fig. S8B), which was associated with reduced responsiveness to EGF and delayed growth (Fig. 4, I and J, and fig. S8C).

Solute transport may be involved in the shrinkage and tubulation of other organelles. In this regard, lysosomes are known to undergo swelling in cells treated with PIKfyve antagonists (*18*). Impaired solute extrusion could account for the volume gain, which is compounded by ongoing membrane fusion that is not compensated for by shrinkage-dependent tubulation and/or vesiculation and scission. Indeed, lysosomes swollen

by inhibition of PIKfyve recovered their ability to tubulate when exposed to hypertonic solution (movie S7) and on removal of the PIKfyve inhibitor but failed to do so when tetrandrine was present (fig. S9). Overexpression of TPC2 alone caused lysosomes to become more tubular, suggesting that the channel may be involved in this process (fig. S10 and movie S8).

We propose a role for ion fluxes in the endocytic pathway: Extrusion of osmotically abundant ions and solutes, accompanied by water, cause vacuolar and vesicular shrinkage leading to crenation of the membrane, forming convex protrusions that stabilize proteins with BAR or similar curvature-sensing domains. These proteins, in turn, foster tubulation and eventual scission that is critical for interorganellar traffic. In macrophages that continuously survey the extracellular space, an ongoing ion efflux from the endocytic pathway drives the resolution of fluid that is taken up during this process, supporting recycling of receptors to maintain their function. Thus, the resolution of organelles formed during surveillance is necessary to preserve tissue integrity.

REFERENCES AND NOTES

1. R. M. Steinman, S. E. Brodie, Z. A. Cohn, *J. Cell Biol.* **68**, 665–687 (1976).
2. J. A. Swanson, *Nat. Rev. Mol. Cell Biol.* **9**, 639–649 (2008).
3. J. A. Swanson, C. Watts, *Trends Cell Biol.* **5**, 424–428 (1995).
4. S. Yoshida, A. D. Hoppe, N. Araki, J. A. Swanson, *J. Cell Sci.* **122**, 3250–3261 (2009).
5. S. Uderhardt, A. J. Martins, J. S. Tsang, T. Lämmermann, R. N. Germain, *Cell* **177**, 541–555.e17 (2019).
6. Y. Sakurai *et al*., *Science* **347**, 995–998 (2015).
7. J. E. Lattin *et al*., *Immunome Res.* **4**, 5 (2008).
8. J. She *et al*., *Nature* **556**, 130–134 (2018).
9. X. Wang *et al*., *Cell* **151**, 372–383 (2012).
10. S. K. Dove *et al*., *Nature* **390**, 187–192 (1997).
11. X. Li *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 21165–21170 (2013).
12. B. J. Peter *et al*., *Science* **303**, 495–499 (2004).
13. M. C. Kerr *et al*., *J. Cell Sci.* **119**, 3967–3980 (2006).
14. M. L. Dustin, *Microbiol. Spectr.* **4**, MCHD-0034-2016 (2016).
15. J. Castonguay *et al*., *Sci. Rep.* **7**, 10038 (2017).
16. C. Grimm *et al*., *Nat. Commun.* **5**, 4699 (2014).
17. C. Commisso *et al*., *Nature* **497**, 633–637 (2013).
18. H. B. Jefferies *et al*., *EMBO Rep.* **9**, 164–170 (2008).

ACKNOWLEDGMENTS

We thank K. Aranda, Z. Liu, and M. Capurro for tissue preparation and genotyping. We are grateful to the Transgenic and Chimeric Mouse Core of the University of Pennsylvania for the generation of the TPC1 line. **Funding:** This work was supported by grant FDN-143202 from CIHR to S.G., the Intramural Research Program of the NIH to R.N.G. and S.U., and by NIH grants R01 HL147379 and R01 GM133172 to D.R. C.M.B. was supported by the Wellcome Trust Sir Henry Wellcome Postdoctoral Fellowship. **Author contributions:** S.A.F. and S.G. conceptualized the project. S.A.F., S.U., A.S., and S.G. designed the research and wrote the paper. S.A.F., S.U., A.S., R.F.C., C.M.B., S.M., P.B., and J.P. performed the experiments and analyzed the results. R.N.G. and D.R. devised the methodology, provided resources, and edited the manuscript. **Competing interests:** None declared. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/301/suppl/DC1
Materials and Methods
Figs. S1 to S10
References (*19*–*28*)
Movies S1 to S8

View/request a protocol for this paper from *Bio-protocol*.

8 February 2019; resubmitted 8 October 2019
Accepted 25 November 2019
Published online 5 December 2019
10.1126/science.aaw9544

MAMMALIAN EVOLUTION

Integrated hearing and chewing modules decoupled in a Cretaceous stem therian mammal

Fangyuan Mao[1,2,3]*, Yaoming Hu[1]†, Chuankui Li[1,2]†, Yuanqing Wang[1,2], Morgan Hill Chase[4], Andrew K. Smith[4], Jin Meng[3,5]*

On the basis of multiple skeletal specimens from Liaoning, China, we report a new genus and species of Cretaceous stem therian mammal that displays decoupling of hearing and chewing apparatuses and functions. The auditory bones, including the surangular, have no bone contact with the ossified Meckel's cartilage; the latter is loosely lodged on the medial rear of the dentary. This configuration probably represents the initial morphological stage of the definitive mammalian middle ear. Evidence shows that hearing and chewing apparatuses have evolved in a modular fashion. Starting as an integrated complex in non-mammaliaform cynodonts, the two modules, regulated by similar developmental and genetic mechanisms, eventually decoupled during the evolution of mammals, allowing further improvement for more efficient hearing and mastication.

In non-mammaliaform cynodonts, the primary jaw joint served for both chewing (mastication) and hearing (sound transfer) functions. In mammals, the two functions and related structures are separated, characterized by a single-boned lower jaw and a tri-ossicular middle ear. Although the primary jaw joint and postdentary bones differ from the mammalian auditory bones in morphology, their homologies have been demonstrated by developmental, genetic, and paleontological evidence (*1*–*7*). Here, we report a new genus and species of symmetrodont mammal from the Early Cretaceous Jehol Biota, China. The unprecedented preservation of the specimens displays key structures related to hearing and chewing morphologies, such as tooth crown structures, ossified Meckel's cartilage and its lodging groove on the dentary, and the auditory bones (the stapes, malleus, incus, ectotympanic, and surangular) (Figs. 1 to 3). The configuration of the auditory bones most probably represents the beginning stage of the definitive mammalian middle ear (*1*) and narrows the morphological gap between the former and the transitional mammalian middle ear (*8*). Given their homologies and similar developmental generic patterning mechanisms (*2*–*6*, *9*), the hearing and chewing apparatuses are hypothesized as two integrated modules that were decoupled during mammalian evolution. The final disassociation of the two modules could have increased the capacity to generate heritable phenotypic variations (*10*) and thus provided the potential for improvement of hearing and chewing functions in future therians.

Origolestes lii, gen. et sp. nov. (*11*) has acute-angled molars that are typical for spalacotheriid "symmetrodontans" (*11*–*17*), an extinct group of stem therian mammals (*18*) (Figs. 1 and 2). The embrasure between upper molars is a narrow but transversely deep wedge-shaped space, corresponding to a narrow cusp a of the lower molar. In contrast, the lower molar is transversely narrower but mesiodistally long so that the embrasure between lower molars is open and shallow, being able to accommodate a broad cusp A of the upper molar. Palatal fossae on the palate exist lingual to upper molars (fig. S3) and must have received tall cusps of lower molars while the lower jaws were at rest position in life. During the evolution of the tribosphenic tooth pattern, the upper molar was lingually extended by addition of the protocone and the lower molar developed the talonid; these changes let the protocone bite in the talonid for grinding. Additionally, the lower molar cusps no longer rested by biting in the palatal fossae, which could help to protect the palate and gum as well as increase the space of the mouth cavity available for food holding and processing.

Because the lower teeth were positioned lingual to the uppers at rest position, during mastication the opened lower jaw would close dorsolabially with a degree of eversion so that cusp a of the lower molar can fit in the narrow embrasure between upper molars (Fig. 2). Then, the lower molar would move dorsolingually in a curved path during the power stroke because the functional surfaces of the wedge-shaped tooth cusps are convex. Thus, in addition to the transverse component of the jaw movement, the tooth shape and occlusal relation dictate that the mandible had to invert by rolling inward relative to its long axis during jaw closing, and the unfused symphysis (Fig. 3) allows such eversion and inversion of the lower jaw, similar to other mammaliaforms (*19*). These features imply that components of both jaw yaw (*20*) and rolling (*19*) existed during mastication of *Origolestes*. The acute-angled molars of *Origolestes* are relatively wider than the triconodont molars but narrower than the tribosphenic ones; the degree of the two movements would be intermediate between those with triconodont and tribosphenic molars. Although jaw yaw and rolling may be primitive mammaliaform features (*19*, *20*), they probably played a role in the decoupling of the auditory bones from the dentary and eventually from the Meckel's cartilage during mammalian evolution.

The long ossified Meckel's cartilage is rod-like, broad posteriorly but tapering anteriorly, with its posterior end bending medially (Fig. 3 and fig. S6). The stapes has a large process for insertion of the stapedius muscle (Fig. 3 and figs. S6 and S7). The incus articulates with the malleus and possibly the surangular and anchors in the epitympanic recess by a dorsal plate; its stapedial process curves medially to articulate the stapes. The malleus has a short anterior process and a blunt manubrium, a neomorphic structure in mammals (*2*, *3*, *21*). The surangular is present as a distinct bone dorsolateral to the malleus body. In some non-mammaliaform cynodonts, a surangular boss is dorsolateral to the primary jaw joint and may have functioned to reduce the compressive load borne by the quadrate in life (*1*); a similar structure was interpreted in the same position in *Liaoconodon* (*8*) (Fig. 3). The surangular has been reported in the euharamiyidan *Arboroharamiya* but remains poorly known, or unknown, in other Mesozoic mammaliaforms; it may exist in some extant mammals as the accessory malleus (*22*). The ectotympanic is therian-like but has a slim and short ventral limb (= reflected lamina). The malleus, surangular, and ectotympanic are tightly connected, with the thin ectotympanic partly wrapping around the other two elements, so that they likely functioned as one unit to transmit sound vibrations. A gap is between the auditory bones and the distal end of the Meckel's cartilage, probably left by a ligament in life. The lack of bone contact between the two units contrasts with the bone-contact condition of the transitional mammalian middle ear in *Liaoconodon* (Fig. 3).

A sizable stapedius muscle may be inferred from the distinct process for insertion of the stapedius muscle of *Origolestes*, contrasting to the minuscule process on the stapes of extant therians. Moreover, we further postulated that the tensor tympani had inserted to the concavity on the medial side of the malleus body, near the base of the manubrium; a similar condition is present in the malleus of *Liaoconodon*

[1]Key Laboratory of Vertebrate Evolution and Human Origins, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China. [2]CAS Center for Excellence in Life and Paleoenvironment, Beijing 100044, China. [3]Division of Paleontology, American Museum of Natural History, New York, NY 10024, USA. [4]Microscopy and Imaging Facility, American Museum of Natural History, New York, NY 10024, USA. [5]Earth and Environmental Sciences, Graduate Center, City University of New York, New York, NY 10016, USA.
*Corresponding author. Email: jmeng@amnh.org (J.M.); maofangyuan@ivpp.ac.cn (F.M.) †Deceased.

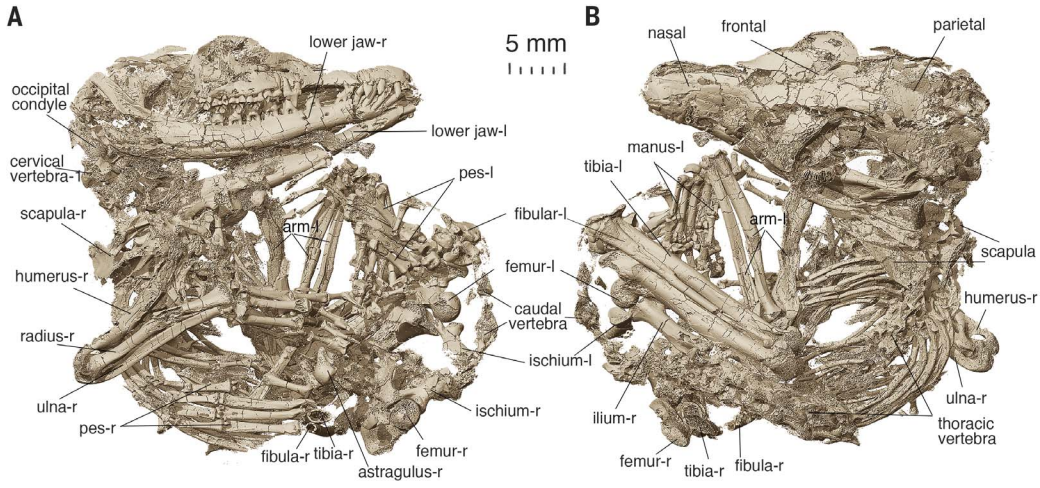


Fig. 1. Skeleton of V14383-1, the holotype specimen of *Origolestes*. (**A** and **B**) The skeleton in roughly ventral and dorsal views. See also figs. S1 and S2.

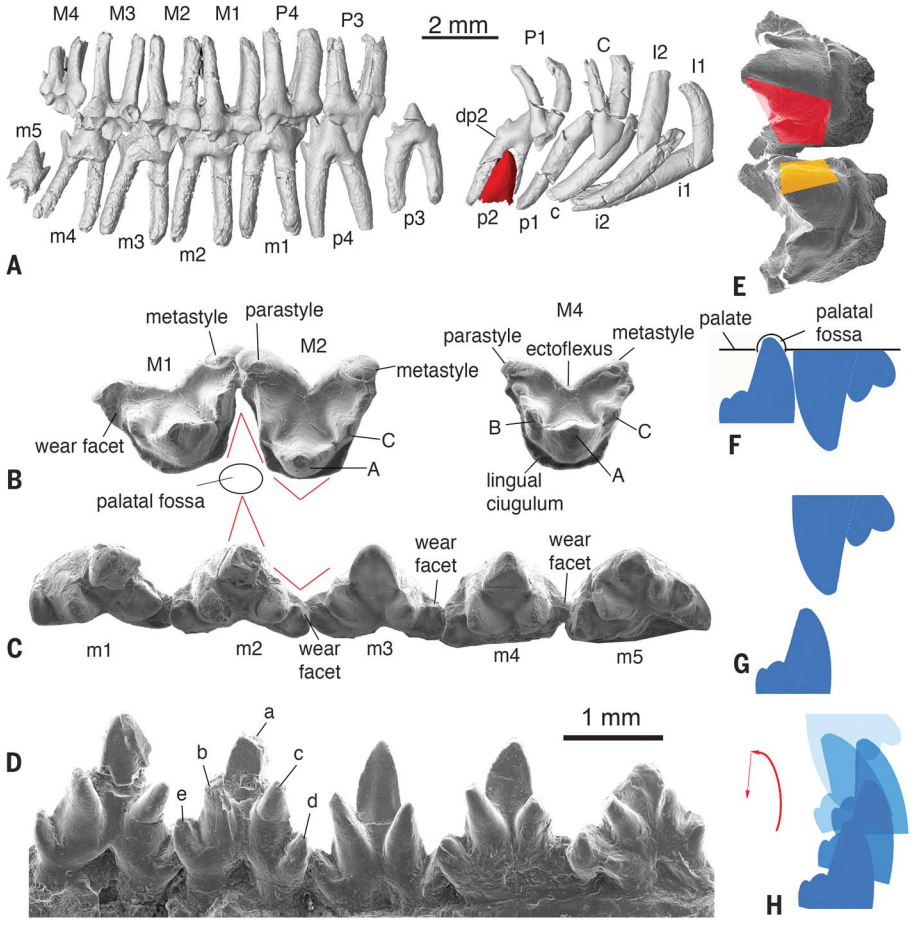


Fig. 2. Dental morphology of *Origolestes*. (**A**) Labial view of right dentitions (V14383-1, holotype). (**B**) Occlusal views of M1-M2 and M4 (V13604). (**C** and **D**) Crown and labial views of left lower molars (V13604). (**E**) Lingual view of M1-M2 (V13604) showing the wedge-shaped embrasure (shearing facets marked by yellow and red). (**F** to **H**) Diagram showing tooth occlusal relationships (at rest, open, and power stroke; red arrow indicates the path). See also figs. S3 to S5.

(Fig. 3). The homologies of the two middle ear muscles have been treated in many studies (*23*–*25*). The tensor tympani was a derivative of the first arch and innervated by the trigeminal nerve, whereas the stapedius is a second-arch derivative innervated by the facial nerve (*2*). Although different interpretations have arisen, it is generally accepted that the tensor tympani (and tensor veli palatini) is derived from the pterygomandibularis (*23*, *25*) and that the stapedius was derived from the levator hyoideus/depressor mandibulae (*24*). These muscles were associated with mastication in non-mammalian tetrapods but transformed into the middle ear of mammals for hearing. Contractions of the muscles dampen sound-induced oscillations of tympanum and middle ear bones and reduce sound amplitude, thus protecting the inner ear from intense sound signals (*9*, *26*).

In non-mammaliaform cynodonts, such as *Thrinaxodon*, the postdentary bones had functioned for jaw articulation and sound transfer to the inner ear (*1*); even the stapes had played a role in mastication to resist medial displacement of the quadrate during chewing (*27*). Thus, the hearing and chewing apparatuses formed a structurally and functionally integrated complex. In the mandibular middle ear of *Morganucodon*, the postdentary bones were greatly reduced and the secondary jaw articulation was formed. In the transitional mammalian middle ear of *Liaoconodon*, the auditory bones were detached from the dentary but retained substantial bone contact with the Meckel's cartilage, so that hearing and chewing functions would have still interfered with each other. In *Origolestes*, such a bone contact is lost, showing the decoupling configuration of the hearing and chewing apparatuses that had been predicted in previous

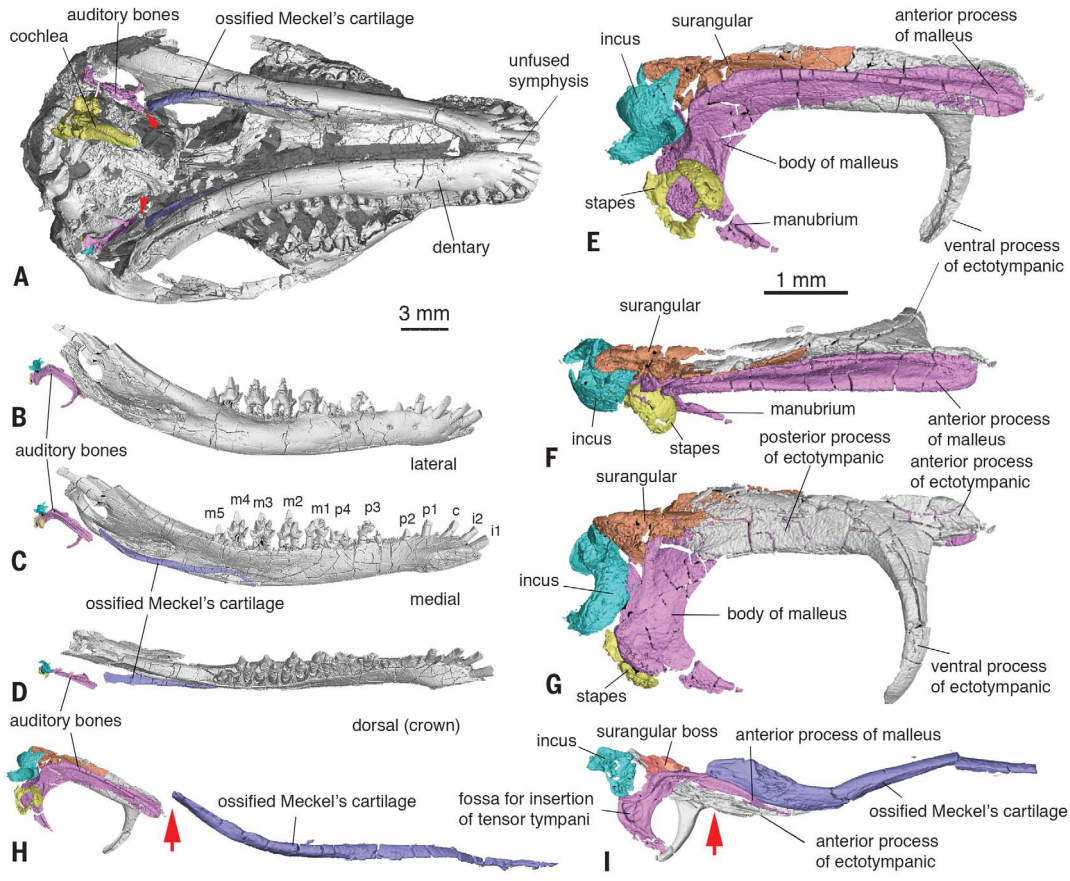


Fig. 3. Skull, lower jaws, and auditory bones of *Origolestes*. (**A**) Ventral view of the skull (JZD-DB0064), showing the relationship of the lower jaws, ossified Meckel's cartilage, auditory bones, and the right inner ear with the petrosal bone digitally removed (see fig. S8). (**B** to **D**) Lateral, medial, and dorsal views of the left mandible, Meckel's cartilage, and auditory bones. (**E** to **G**) Close-up medial, dorsal, and lateral views of left auditory bones. (B) and (G) are reversed to allow alignment of features. (**H** and **I**) Medial views of auditory bones and Meckel's cartilage of *Origolestes* (H) and *Liaoconodon* (I). Red arrows point to the gap between auditory bones and Meckel's cartilage in *Origolestes* and bone contact in *Liaoconodon*. See also figs. S6 and S7.

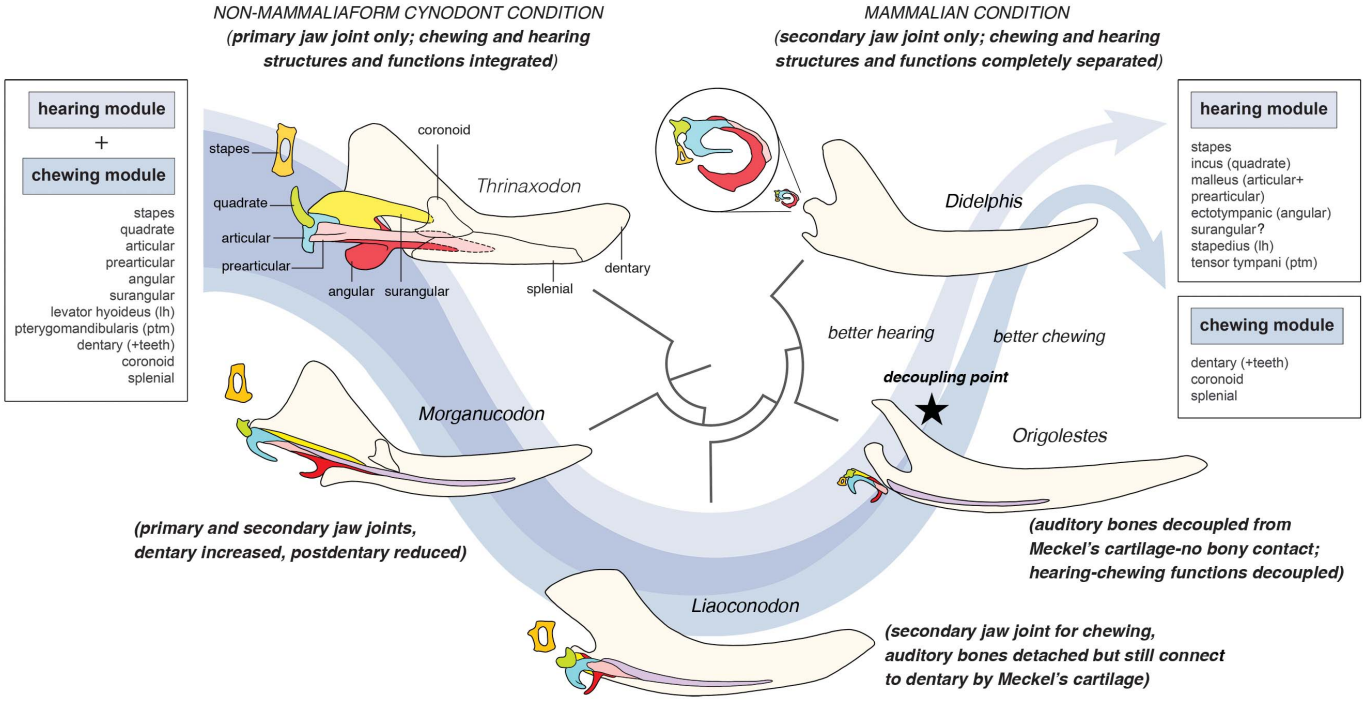


Fig. 4. Diagram illustrating evolutionary stages from the condition in non-mammaliaform cynodonts to that in mammals. See text for discussion. The simplified phylogeny is based on our phylogenetic analysis (fig. S9); most drawings were modified from (*8*).

studies (*1*, *21*). Finally, in the definitive mammalian middle ear, the auditory bones continue to reduce in size and the Meckel's cartilage no longer exists in adults.

Although the postdentary bones in nonmammalian cynodonts are morphologically different from the auditory bones in mammals (Fig. 4), their homologies have been demonstrated by developmental, genetic, and paleontological evidence (*1*–*4*, *6*, *9*) and can be traced back through amniotes (*28*). In particular, it has been shown that genes working in concert in regulating the middle ear bones of mammals also regulate patterning of the jaw joint in non-mammal vertebrates (*5*). Moreover, developmental studies in extant mammals support the findings of the persisting Meckel's cartilage and its groove in Mesozoic mammals (*6*, *9*, *29*). Thus, it is rational to assume that similar genetic regulating mechanisms and developmental pathways existed through the transition from the jaw joint bones to auditory bones during synapsid evolution. Phylogenetically, the definitive mammalian middle ear may have evolved multiple times (fig. S9) (*8*, *21*), with the composition of the ear ossicles in different lineages remaining the same.

On the basis of fossil and developmental genetic evidence, we hypothesize that in synapsids the hearing and chewing apparatuses have evolved as two modules that were regulated by similar genetic and developmental mechanisms, respectively. Starting as a highly integrated structural and functional complex in non-mammaliaform cynodonts, the hearing and chewing modules eventually decoupled, as evidenced in *Origolestes* (Fig. 4), which removed the physical constraint imposed on each other. Such modularity and dissociation would enhance the capacity to generate variation (or evolvability), which may have conferred a selective advantage on modular clades that possessed it (*10*, *30*). The Early Cretaceous *Origolestes* sets a phenotypic and temporal reference that supports the view that during mammalian evolution, the burden-free chewing module could allow modification of jaws, teeth, and their functions for more efficient processing of diverse foods, whereas the hearing module could be further improved for sensitive hearing of high-frequency airborne sounds without being disturbed by mastication (*1*, *8*, *12*, *28*).

REFERENCES AND NOTES

1. E. F. Allin, *J. Morphol.* **147**, 403–437 (1975).
2. M. Takechi, S. Kuratani, *J. Exp. Zool. B* **314**, 417–433 (2010).
3. M. Mallo, *Dev. Biol.* **231**, 410–419 (2001).
4. K. K. Smith, *Dev. Dyn.* **235**, 1181–1193 (2006).
5. A. S. Tucker, R. P. Watson, L. A. Lettice, G. Yamada, R. E. Hill, *Development* **131**, 1235–1245 (2004).
6. N. Anthwal, L. Joshi, A. S. Tucker, *J. Anat.* **222**, 147–160 (2013).
7. W. Maier, I. Ruf, *J. Anat.* **228**, 270–283 (2016).
8. J. Meng, Y. Wang, C. Li, *Nature* **472**, 181–185 (2011).
9. Y. Wang, Y. Zheng, D. Chen, Y. Chen, *Dev. Biol.* **381**, 301–311 (2013).
10. M. Kirschner, J. Gerhart, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8420–8427 (1998).
11. See supplementary materials.
12. A. Crompton, in *Early Mammals*, D. Kermack, K. Kermack, Eds. (Linnean Society of London, 1971), pp. 65–87.
13. R. L. Cifelli, S. K. Madsen, *Geodiversitas* **21**, 167–214 (1999).
14. G. W. Rougier, B. K. Spurlin, P. K. Kik, *Am. Mus. Novit.* **2003**, 1–16 (2003).
15. Z.-X. Luo, P. Chen, G. Li, M. Chen, *Nature* **446**, 288–293 (2007).
16. B. M. Davis, *J. Mamm. Evol.* **18**, 227–244 (2011).
17. J. A. Schultz, T. Martin, *Naturwissenschaften* **101**, 771–781 (2014).
18. T. Harper, G. W. Rougier, *PLOS ONE* **14**, e0209457 (2019).
19. B. S. Bhullar *et al*., *Nature* **566**, 528–532 (2019).
20. D. M. Grossnickle, *Sci. Rep.* **7**, 45094 (2017).
21. E. F. Allin, J. A. Hopson, in *The Evolutionary Biology of Hearing*, D. B. Webster, A. N. Popper, R. R. Fay, Eds. (Springer, 1992), pp. 587–614.
22. O. W. Henson Jr., in *The Handbook of Sensory Physiology: The Auditory System V/I*, W. D. Keidel, W. D. Neff, Eds. (Springer-Verlag, 1974), pp. 39–110.
23. H. R. Barghusen, in *The Ecology and Biology of Mammal-Like Reptiles*, I. N. Hotton, P. D. MacLean, J. J. Roth, E. C. Roth, Eds. (Smithsonian Institution Press, 1986), pp. 253–262.
24. R. Diogo, V. Abdala, N. Lonergan, B. A. Wood, *J. Anat.* **213**, 391–424 (2008).
25. J. M. Ziermann, R. E. Diaz, R. Diogo, *Heads, Jaws, and Muscles: Anatomical, Functional, and Developmental Diversity in Chordate Evolution* (Springer, 2019).
26. S. Mukerji, A. M. Windsor, D. J. Lee, *Trends Amplif.* **14**, 170–191 (2010).
27. K. A. Kermack, F. Mussett, H. W. Rigney, *Zool. J. Linn. Soc.* **71**, 1–158 (1981).
28. G. A. Manley, U. J. Sienknecht, in *The Middle Ear: Science, Otosurgery, and Technology*, S. Puria, R. Fay, A. Popper, Eds. (Springer, 2013), pp. 7–30.
29. N. Anthwal, D. J. Urban, Z.-X. Luo, K. E. Sears, A. S. Tucker, *Nat. Ecol. Evol.* **1**, 0093 (2017).
30. S. B. Carroll, *Nature* **409**, 1102–1109 (2001).

ACKNOWLEDGMENTS

We thank H. Li, J. Li, Z. Gao, and X. Ding (Jizantang Paleontological Museum, Chaoyang City, Liaoning Province, China) for providing the studied specimens housed in their collections; S. Xie (Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing-IVPP) for specimen preparation; Y. Hou, P. Yin (IVPP), and S. Hou (Yinghua Inspection and Testing Co. Ltd., Shanghai) for CT scanning of the specimens; N. Wong (American Museum of Natural History) and A. Shi (IVPP) for drawing and designing the figures; Z. Zhou and X. Wang (IVPP) for discussions of localities and stratigraphies; and two anonymous reviewers for constructive comments. **Funding:** Supported by the National Natural Science Foundation of China (41688103, 41404022); the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (XDB26000000); the Youth Innovation Promotion Association CAS (2019076); and the Kalbfleisch Fellowship, Richard Gilder Graduate School, American Museum of Natural History. **Author contributions:** F.M. and J.M. conceived the study and wrote the paper; F.M., M.H.C., and A.K.S. conducted the CT work; Y.H., C.L., and Y.W. initiated the work and curated the specimens; and all authors (except for Y.H.) edited and approved the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The specimens are accessioned in IVPP and JZT, China; all data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/305/suppl/DC1
Materials and Methods
Systematic Paleontology
Supplementary Text
Figs. S1 to S9
Table S1
References (*31*–*102*)

29 July 2019; accepted 21 November 2019
Published online 5 December 2019
10.1126/science.aay9220

THERMAL CONDUCTIVITY

Phonon hydrodynamics and ultrahigh–room-temperature thermal conductivity in thin graphite

Yo Machida[1]*, Nayuta Matsumoto[1], Takayuki Isono[1], Kamran Behnia[2]*

Allotropes of carbon, such as diamond and graphene, are among the best conductors of heat. We monitored the evolution of thermal conductivity in thin graphite as a function of temperature and thickness and found an intimate link between high conductivity, thickness, and phonon hydrodynamics. The room-temperature in-plane thermal conductivity of 8.5-micrometer-thick graphite was 4300 watts per meter-kelvin—a value well above that for diamond and slightly larger than in isotopically purified graphene. Warming enhances thermal diffusivity across a wide temperature range, supporting partially hydrodynamic phonon flow. The enhancement of thermal conductivity that we observed with decreasing thickness points to a correlation between the out-of-plane momentum of phonons and the fraction of momentum-relaxing collisions. We argue that this is due to the extreme phonon dispersion anisotropy in graphite.

Heat travels in insulators because of the propagation of collective vibrational states of the crystal lattice called phonons. The standard description of this transport phenomenon invokes quasiparticles losing their momentum to the underlying lattice because of collisions along their trajectory (*1*). Gurzhi proposed decades ago that phonons in insulators and electrons in metals can flow hydrodynamically if momentum-conserving collisions among carriers become abundant (*2*). Recently, hydrodynamic regimes for electrons (*3*–*5*) and for phonons (*6*–*10*) have become a subject of renewed attention, partially driven by the aim of quantifying the quasiparticle viscosity.

Unlike particles in an ideal gas of molecules, the phonon momentum is not conserved in all collisions. When scattering between two phonons produces a wave vector exceeding the unit vector of the reciprocal lattice, the excess of momentum is lost to the underlying lattice. These are called Umklapp (U) scattering events, and they require sufficiently large wave vectors. Because cooling reduces the typical wavelength of thermally excited phonons, U scattering rarefies with decreasing temperature, and most collisions among phonons conserve momentum, becoming normal (N) scattering events. In this context, a regime of phonon hydrodynamics emerges that is sandwiched between diffusive and ballistic regimes (*2*). Observations of the hydrodynamic regime include several solids (*8*, *9*, *11*–*14*). In this narrow temperature window, warming multiplies normal collisions, and this enhances the ratio of thermal conductivity to specific heat—called the thermal diffusivity. Observations of this behavior tend to be at cryogenic temperatures.

The domination of N events over U events across a very broad temperature range in graphene led two groups to propose that phonon hydrodynamics might be observed at temperatures outside the cryogenic range (*6*, *7*). However, heat transport measurements in graphene (*15*) are challenging to study by using the standard four-probe steady-state technique. Evidence for second sound, a manifestation of phonon hydrodynamics, was recently found at temperatures exceeding 100 K in graphite (*10*). These observations were in agreement with theoretical expectations (*16*).

The two-dimensional lattice of graphite (Fig. 1A, inset) consists of strong interlayer sp^2 covalent bonds combined with weak intralayer van der Waals bonds. The strength of the in-plane and the out-of-plane couplings differs by two orders of magnitude. This dichotomy makes graphite easily cleavable down to the single-layer graphene form (*17*). The bonding of graphite also creates two distinct Debye temperatures, one for the in-plane and the other for the out-of-plane atomic vibrations (*18*). This induces a large difference between in-plane and out-of-plane thermal conductivities (*19*). The experimentally measured thermal conductivity (*19*–*23*) shows a roughly similar temperature dependence. However, there is a large variety in the reported magnitude of in-plane thermal conductivity, which at room temperature can vary between 72 and 2100 W/m·K (*19*), a feature attributed to the unavoidable presence of the stacking faults and contamination of the in-plane data by a contribution from *c*-axis flow. As we will see below, new insight is provided by a thickness-dependent study on the same sample.

We measured the in-plane thermal conductivity (κ) of commercially available highly oriented pyrolytic graphite (HOPG) samples, all peeled from a thick mother sample, with a standard steady-state one-heater–two-thermometers technique in high vacuum

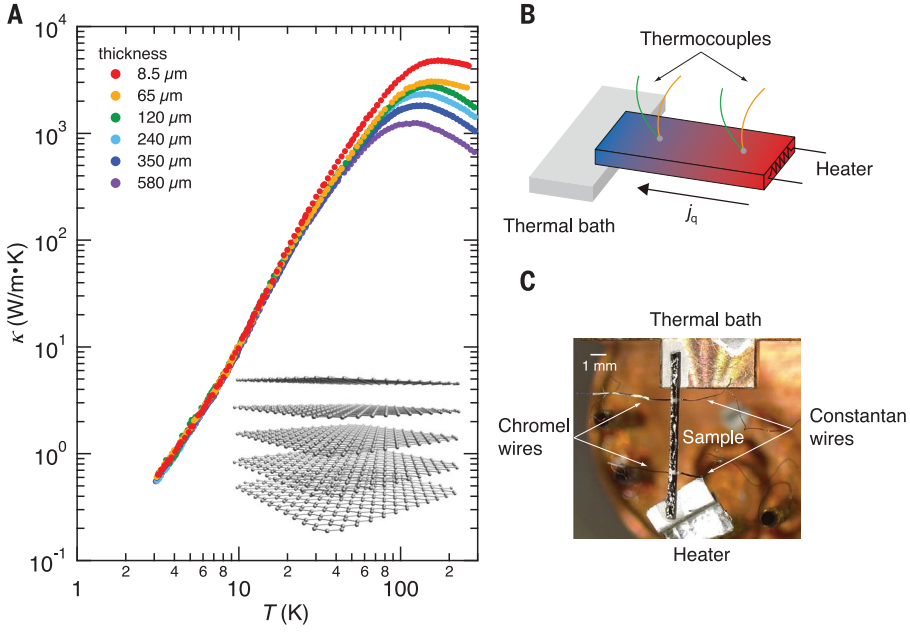


Fig. 1. Thermal conductivity and experimental setup. (**A**) Temperature dependence of in-plane thermal conductivity of graphite with thicknesses ranging from 580 to 8.5 μm on a logarithmic scale. Inset shows side view of the crystal structure of graphite. A schematic illustration (**B**) and a photo (**C**) of the measurement setup for the thermal conductivity. Heat current j_q generated by a heater on one end of the sample passes through the sample toward the thermal bath. Temperature difference developed in the sample is determined by two pairs of thermocouples.

[1]Department of Physics, Gakushuin University, Tokyo 171-8588, Japan. [2]Laboratoire Physique et Etude de Matériaux (CNRS-Sorbonne Université-ESPCI), PSL Research University, 75005 Paris, France.
*Corresponding author. Email: yo.machida@gakushuin.ac.jp (Y.M.); kamran.behnia@espci.fr (K.B.)

(Fig. 1). We tested the reliability by measuring the thermal conductivity of a long, thin silver foil with a thermal resistance comparable to that of our most thermally resistive sample and quantifying the small deviation from the Wiedemann-Franz law (*24*). For samples with thicknesses ranging from 8.5 to 580 μm, we found identical κ behavior below 20 K and a steady thickness evolution for κ with increasing temperature above 20 K.

We compared the temperature dependence of κ in the thickest sample (580 μm) with the measured specific heat (Fig. 2A). We found that κ peaks around 100 K, similar to other measurements (*20*, *21*, *23*). Below this maximum, κ quickly decreases and roughly follows a $T^{2.5}$ dependence, close to the specific heat trend below 10 K (*25*). The specific heat (C) temperature behavior deviated from the T^3 expected from the Debye approximation. However, this behavior is not strictly observed in most real solids, owing to unequal distribution of phonon weights. The 2.5 exponent has been attributed in graphite to an admixture of T^3 and T^2 contributions by out-of-plane and in-plane phonons, respectively (*26*). This unusual exponent may have obscured the Poiseuille regime, which is usually associated with faster-than-cubic thermal conductivity (*2*).

Closer examination of the parallel evolution of thermal conductivity and specific heat can help unveil the Poiseuille regime as κ evolves faster than C above 10 K and slower below 10 K (Fig. 2A). Plotting $\kappa/T^{2.5}$ and $C/T^{2.5}$ makes this difference easier to recognize (Fig. 2B). Upon warming, $\kappa/T^{2.5}$ shows a pronounced maximum above 10 K, whereas $C/T^{2.5}$ gradually decreases. The thermal diffusivity, D_{th}, is the ratio of thermal conductivity to specific heat (expressed in proper units of J/K mol). We found that D_{th} has a nonmonotonic temperature dependence between 10 and 20 K (Fig. 2C). The phonon hydrodynamic picture provides a straightforward interpretation of this feature. Warming leads to enhanced momentum exchange among phonons, because the fraction of collisions that conserve momentum increases. As a consequence, heat diffusivity rises. If all phonons had the same mean free path irrespective of their branch and wave vector, this would also imply a rise in the effective mean free path. The Poiseuille maximum around 40 K and the Knudsen minimum around 10 K, where diffuse boundary scattering rate is effectively increased because of N scattering, define the boundaries of this hydrodynamic window.

We found that the electron contribution is negligibly small in the temperature range of interest by determining the Lorenz ratio (L/L_0). We measured electrical conductivity, σ, to quantify $L = \frac{\kappa}{\sigma T}$ and compare it with $L_0 = 2.44 \times 10^{-8}$ W-ohm/K^2. This results in a ratio between 100 and 1000 above the Knudsen minimum (Fig. 2D).

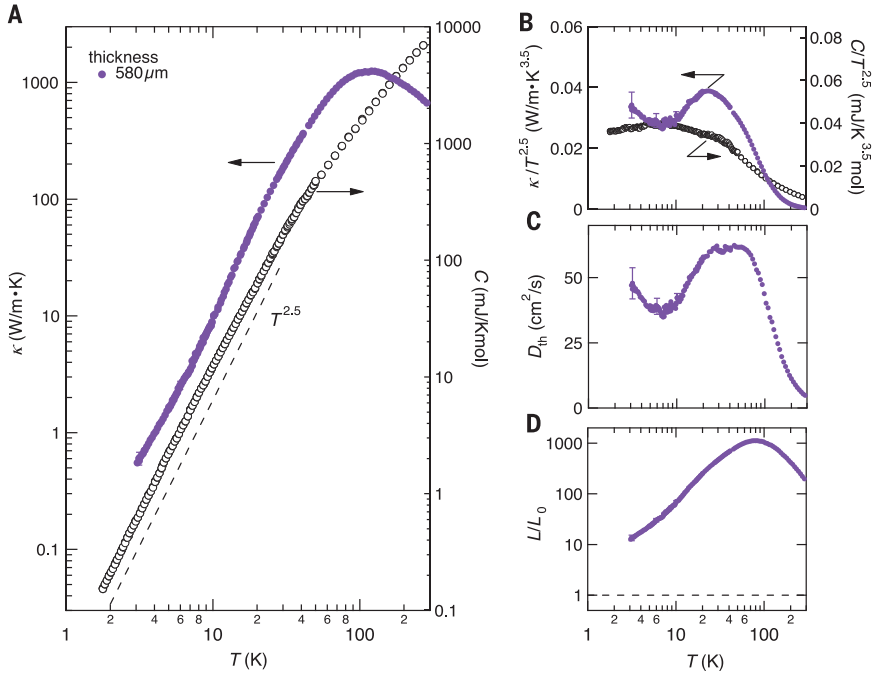


Fig. 2. Hydrodynamic heat transport. (**A**) Temperature dependence of in-plane thermal conductivity κ (left axis) and specific heat C (right axis) of the 580-μm-thick graphite sample. (**B**) κ divided by $T^{2.5}$ (left axis) and C divided by $T^{2.5}$ (right axis) as a function of temperature. A pronounced maximum is seen only in $\kappa/T^{2.5}$ above 10 K. This yields a maximum in temperature dependence of thermal diffusivity D_{th} (**C**). Dominant phonon contribution in κ is indicated by a large Lorenz ratio L/L_0 shown in (**D**).

The behavior that we observed for κ and C is not due to outstanding sample quality. Comparable features can be found in published data (*20*, *21*, *24*) but appear to have gone unnoticed. Our mother sample was an average HOPG containing both stable isotopes of carbon, (~99% ^{12}C, ~1% ^{13}C). Our results support the conjecture that phonon hydrodynamics can occur without isotopic purity (*8*).

We measured an increased κ as we decreased sample thickness (Fig. 3A). We performed successive measurements after changing the thickness (t) of the sample along the c axis, maintaining the sample width (w = 350 μm) and the distance (l) between contacts for the thermal gradient to be long enough compared to the thickness ($l/t > 10$) (*24*). The trend is the opposite of observations for black phosphorus (*8*). With respect to the hydrodynamic regime, thinning leads to an amplification of the nonmonotonic behavior of thermal diffusivity. This drives the Poiseuille peak to become sharper and toward higher temperatures. Eventually, D_{th} of the thinnest sample shows a sharp maximum at 100 K. Second sound in graphite was observed near this temperature (*10*). The thickness dependence vanishes below 10 K, presumably because the phonon mean free path in this range is set by the average crystallite size (*19*), which does not depend on thickness. Another possible origin of the thickness-independent low-temperature thermal conductivity is an intrinsic scattering of phonons by mobile electrons.

The thermal conductivity in our 240-μm-thick sample is in reasonable agreement with previous observations on a similar thickness graphite (*22*). The in-plane κ that we measured for the 8.5-μm-thick sample was ~4300 W/m·K. This exceeds the value for an isotopically pure graphene sample (*27*) and is higher than that of other bulk solids. The value is twice that of natural abundance diamond (*28*) and about three times larger than high-purity crystalline BAs (*29*–*31*). At room temperature, reducing the thickness by two orders of magnitude leads to a fivefold increase in κ (Fig. 3C). Although the κ that we measured is already comparable with the highest values reported in single-layer graphene (κ ≈ 3000 to 5000 W/m·K) (*27*, *32*), our data do not saturate in the low-thickness limit. In contrast to suspended graphene over a trench of 3 μm (*32*), our samples are millimetric in length. Given the quasi-ballistic trajectory of phonons, we make the reasonable assumption that in-plane dimensions matter in setting the amplitude of thermal conductivity. This would imply that the ceiling is higher

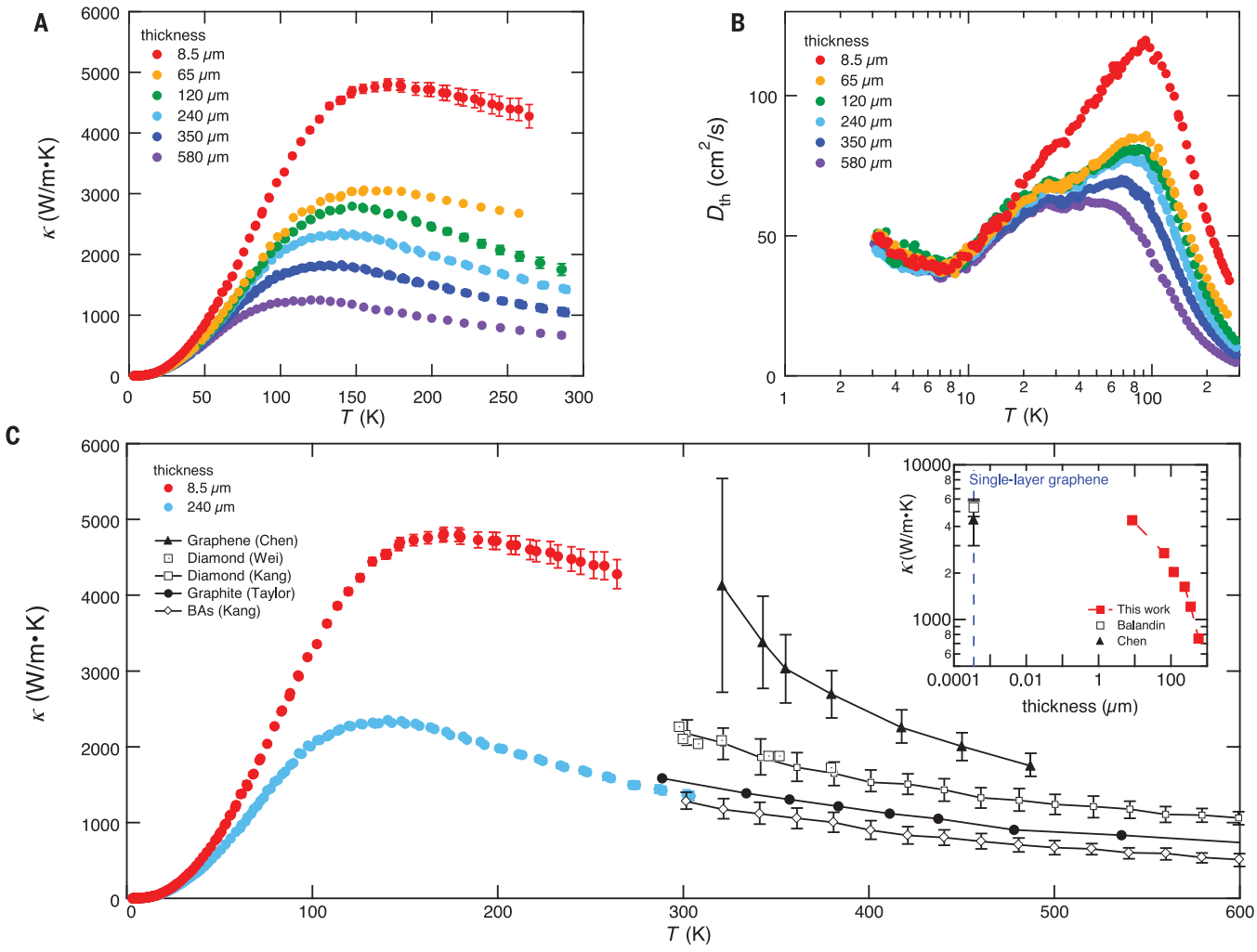


Fig. 3. Thickness dependence of thermal conductivity. (**A**) Temperature dependence of in-plane thermal conductivity κ for various sample thicknesses. In the thinnest sample, κ attains the largest value (~4300 W/m·K) known in any bulk system near room temperature. (**B**) Temperature dependence of thermal diffusivity D_{th} for various sample thicknesses. The maximum in D_{th} forms a sharp, single peak with decreasing thickness. (**C**) Our data are compared with those of ultrahigh–thermal conductivity materials (*22*, *27*–*29*). The inset shows thickness dependence of thermal conductivity at 250 K. κ of the thinnest sample is comparable with the high values reported in single-layer graphene (*27*, *32*).

than previously believed, and thinner samples with larger aspect ratio should display even larger conductivity. Although several theoretical works have predicted a robust hydrodynamic regime in graphene (*6*, *7*) and its persistence in graphite (*16*), none examined the issue of thickness dependence.

To try to understand the origin of our observation, we scrutinized the occurrence of U and N collisions, given the phonon dispersion of graphite (*15*, *33*) (Fig. 4). We show the calculation of Nihira and Iwata (*33*) from a semicontinuum model for the in-plane and out-of-plane dispersion of longitudinal (LA), in-plane transverse (TA), and out-of-plane transverse (ZA) acoustic phonons along the ΓM and ΓA directions (Fig. 4B). The model parameters (velocities and elastic constants) were determined by using the best account of experimental specific heat data from 0.5 to 500 K (*33*). The two orientations show a marked contrast regarding the typical wavelength of thermally excited phonons and requirements for U scattering. At 300 K (or 200 cm^{-1}), the typical in-plane wave vector of the LA mode is only 0.1 of the Brillouin zone (BZ) width. This makes U collisions extremely rare (Fig. 4C), because to create a phonon with a wave vector larger than half of the BZ width, the average wave vector of each colliding phonon needs to be 0.25 of the BZ width. The fundamental reason behind the scarcity of U collisions and the emergence of hydrodynamics resides within this simple feature. The situation is radically different for out-of-plane wave vectors. Even at 50 cm^{-1}, a thermally excited phonon can have an out-of-plane wave vector that is one-fourth of the BZ height. Above 90 cm^{-1} (corresponding to 130 K), out-of-plane phonons are all thermally excited (*33*), and their peak wavelength is half of the BZ height. Any additional momentum along this orientation can "kick" them out of the BZ. A small *c*-axis component in the momentum exchanged by colliding phonons suffices for the collision to become a U event (Fig. 4D) and the heat flow to degrade.

Our observation implies a reduction in the relative weight of U collisions as the sample is thinned, because attenuating the relative rate of U collisions would extend the hydrodynamic window and enhance thermal conductivity. We note that the spacing between discrete available states in the reciprocal space depends on thickness. Therefore, the total number of states with out-of-plane momentum is inversely proportional to the thickness. It

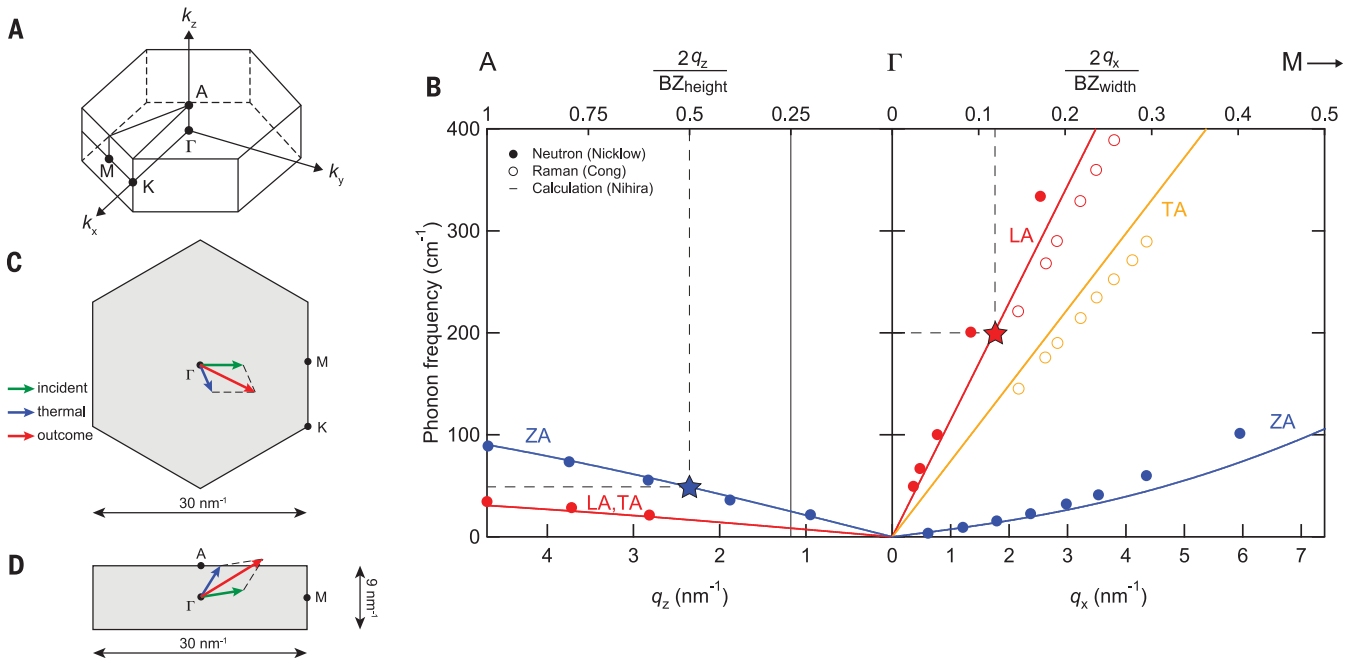


Fig. 4. Phonon dispersions. (**A**) First Brillouin zone (BZ) of graphite. (**B**) Calculated dispersions of acoustic phonon blanches along the ΓA and ΓM directions of BZ (*33*), together with the experimental data obtained by neutron (*34*) and Raman scattering (*35*). BZ in the ΓKM plane (**C**) and ΓMA plane (**D**). Collision between the in-plane component of an incident phonon (green arrow) and a thermally excited phonon (blue arrow) remains N, because the in-plane wave vector of the thermal phonon is only a small fraction of the BZ width even at 300 K (or 200 cm^{-1}). Hence, the wave vector of the outcome phonon (red arrow) does not exceed one-half of the BZ width. By contrast, the out-of-plane wave vector of a thermal phonon is one-fourth of the BZ height for frequencies as low as 50 cm^{-1}. Therefore, the collision becomes U, if the in-plane traveling phonon happens to possess a small out-of-plane component.

is true that only a small fraction of the BZ is wiped out by the finite size. However, different collision mechanisms are competing for phase space, and reducing the thickness not only reduces the population of the out-of-plane phonons but also amplifies boundary scattering. Heat-carrying phonons can suffer either a U collision with an out-of-plane phonon or a (more or less) specular collision at the boundary. Thus, reducing the thickness, by substituting a fraction of U collisions with specular boundary reflection, would limit the degradation of the heat flow.

A satisfactory account of thickness dependence of thermal conductivity in both HOPG and black phosphorus (*8*) is lacking. Scattering at the boundaries and imperfect transmission across interfaces between partially twisted graphene layers require further scrutiny. Serious theoretical calculations are needed to explain our findings.

REFERENCES AND NOTES

1. J. M. Ziman, *Electrons and Phonons: The Theory of Transport Phenomena in Solids* (Oxford Univ. Press, 2001).
2. R. N. Gurzhi, *Sov. Phys. Usp.* **11**, 255–270 (1968).
3. D. A. Bandurin *et al.*, *Science* **351**, 1055–1058 (2016).
4. J. Crossno *et al.*, *Science* **351**, 1058–1061 (2016).
5. P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, A. P. Mackenzie, *Science* **351**, 1061–1064 (2016).
6. S. Lee, D. Broido, K. Esfarjani, G. Chen, *Nat. Commun.* **6**, 6290 (2015).
7. A. Cepellotti *et al.*, *Nat. Commun.* **6**, 6400 (2015).
8. Y. Machida *et al.*, *Sci. Adv.* **4**, eaat3374 (2018).
9. V. Martelli, J. L. Jiménez, M. Continentino, E. Baggio-Saitovitch, K. Behnia, *Phys. Rev. Lett.* **120**, 125901 (2018).
10. S. Huberman *et al.*, *Science* **364**, 375–379 (2019).
11. L. P. Mezhov-Deglin, *Zh. Eksp. Teor. Fiz.* **49**, 66–79 (1965).
12. W. C. Thomlinson, *Phys. Rev. Lett.* **23**, 1330–1332 (1969).
13. V. N. Kopylov, L. P. Mezhov-Deglin, *Zh. Eksp. Teor. Fiz.* **65**, 720–734 (1973).
14. N. N. Zholonko, *Phys. Solid State* **48**, 1678–1680 (2006).
15. D. L. Nika, A. A. Balandin, *Rep. Prog. Phys.* **80**, 036502 (2017).
16. Z. Ding *et al.*, *Nano Lett.* **18**, 638–649 (2018).
17. A. K. Geim, K. S. Novoselov, *Nat. Mater.* **6**, 183–191 (2007).
18. J. A. Krumhansl, H. Brooks, *J. Chem. Phys.* **21**, 1663–1669 (1953).
19. G. N. Slack, *Phys. Rev.* **127**, 694–701 (1962).
20. J. C. Bowman, J. A. Krumhansl, J. T. Meers, *Industrial Carbon and Graphite* (Society of Chemical Industry, London, 1958), p. 52.
21. M. G. Holland, C. A. Klein, W. D. Straub, *J. Phys. Chem. Solids* **27**, 903–906 (1966).
22. R. Taylor, *Philos. Mag.* **13**, 157–166 (1966).
23. D. T. Morelli, C. Uher, *Phys. Rev. B Condens. Matter* **31**, 6721–6725 (1985).
24. Materials and Methods are available as Supplementary Materials on *Science* Online.
25. M. G. Alexander, D. P. Goshorn, D. G. Onn, *Phys. Rev. B Condens. Matter* **22**, 4535–4542 (1980).
26. K. Komatsu, *J. Phys. Soc. Jpn.* **10**, 346–356 (1955).
27. S. Chen *et al.*, *Nat. Mater.* **11**, 203–207 (2012).
28. L. Wei, P. K. Kuo, R. L. Thomas, T. R. Anthony, W. F. Banholzer, *Phys. Rev. Lett.* **70**, 3764–3767 (1993).
29. J. S. Kang, M. Li, H. Wu, H. Nguyen, Y. Hu, *Science* **361**, 575–578 (2018).
30. S. Li *et al.*, *Science* **361**, 579–581 (2018).
31. F. Tian *et al.*, *Science* **361**, 582–585 (2018).
32. A. A. Balandin *et al.*, *Nano Lett.* **8**, 902–907 (2008).
33. T. Nihira, T. Iwata, *Phys. Rev. B Condens. Matter Mater. Phys.* **68**, 134305 (2003).
34. R. Nicklow, N. Wakabayashi, H. G. Smith, *Phys. Rev. B Condens. Matter Mater. Phys.* **5**, 4951–4962 (1972).
35. X. Cong *et al.*, *Carbon* **149**, 19–24 (2019).

ACKNOWLEDGMENTS

We thank S. Kurose for the contribution in the early stage of this study and A. Subedi for discussions. We also thank M. Tsubota and M. Watanabe for technical support. **Funding:** This work was supported by the Japan Society for the Promotion of Science Grant-in-Aids KAKENHI 16K05435, 17KK0088, and 19H01840 and by the Agence Nationale de la Recherche (ANR-18-CE92-0020-01). **Author contributions:** Y.M. and K.B. conceived of and designed the study. Y.M., N.M., and T.I. performed the transport and specific heat measurements. Y.M. and K.B. wrote the manuscript with assistance from all the authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/309/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S7
Table S1
References (*36*–*41*)

10 October 2019; accepted 18 December 2019
10.1126/science.aaz8043

ORGANOMETALLICS

Fluorination of arylboronic esters enabled by bismuth redox catalysis

Oriol Planas*, Feng Wang*, Markus Leutzsch, Josep Cornella†

Bismuth catalysis has traditionally relied on the Lewis acidic properties of the element in a fixed oxidation state. In this paper, we report a series of bismuth complexes that can undergo oxidative addition, reductive elimination, and transmetallation in a manner akin to transition metals. Rational ligand optimization featuring a sulfoximine moiety produced an active catalyst for the fluorination of aryl boronic esters through a bismuth (III)/bismuth (V) redox cycle. Crystallographic characterization of the different bismuth species involved, together with a mechanistic investigation of the carbon-fluorine bond-forming event, identified the crucial features that were combined to implement the full catalytic cycle.

Homogeneous transition-metal catalysis has revolutionized organic synthesis, enabling fast and direct construction of complex functionality. These reactions rely, in large part, on the capacity of noble metals to cycle easily between different oxidation states (Fig. 1A) (*1*). With the goal of providing more sustainable strategies in catalysis, efforts have shifted to unveil the reactivity of more Earth-abundant, first-row transition metals (Fe, Ni, Co, Cu, Mn, and Cr) (*2*–*4*). Chemists have also sought to discover and exploit transition-metal–like reactivity among elements beyond the d-block (*5*–*7*). The concept of the frustrated Lewis pair (FLP) can be applied to boron and phosphorus cooperatively to promote transformations traditionally restricted to transition metals (*8*). Furthermore, methodologies based on alkali (*9*, *10*), alkaline earth (*11*), and group 13 to 17 elements are emerging as acceptable alternatives in certain domains of catalysis (*12*–*16*). These strategies represent useful platforms for chemical synthesis. However, the exploitation of the redox properties of main-group elements in catalysis to access new modes of reactivity is still in its infancy (*17*) and remains a major challenge in organometallic and organic chemistry. We therefore focused our attention on bismuth (Bi), an Earth-abundant and inexpensive main-group element (*18*) with under-explored redox properties (*19*). The use of Bi in catalysis has relied largely on its Lewis acidic properties (*20*), where recent interest has revitalized its use in fields such as C–H activation (*21*), carbonylation (*22*), transfer hydrogenation (*23*), and as the initiator in radical processes (*24*–*26*). We sought to prepare a Bi complex capable of mimicking the canonical, fundamental steps in a transition-metal catalytic cycle: transmetallation (TM), oxidative addition (OA), and reductive elimination (RE) (Fig. 1B). To this end, we focused on the oxidative fluorination of aromatic boronic acids, a transformation that is highly coveted in the pharmaceutical and agrochemical industries (*27*). This reaction is feasible using stoichiometric transition metals, such as Cu (*28*–*30*), Pd (*31*), and Ag (*32*), or hypervalent iodine compounds (*33*). The sole catalytic variant makes use of trifluoroborate aryl salts as substrates, which are converted to aryl fluorides through single-electron transfer processes catalyzed by Pd (*34*). On the basis of the accessibility to Bi compounds in different oxidation states (*35*), we report that the rationally designed bismine complex (**I**) catalyzes fluorination of aryl boronic esters through a Bi(III)/Bi(V) redox cycle.

We hypothesized that a rationally designed ligand would be crucial to exploit the Bi(III)/Bi(V) redox couple. Building on precedents in Bi coordination chemistry (*36*), we sought to take advantage of Bi's capacity for accommodating additional neutral ligands in its coordination sphere, thereby affecting the geometry and the electronics of the Bi center (*37*). Accordingly, we chose a tethered bis-anionic aryl ligand, featuring a linking sulfonyl group in the backbone (Fig. 2A). We predicted that the use of a tether ligand would become important for controlling the geometry in subsequent high-valent intermediates in the catalytic cycle, as Bi(V) compounds are known to undergo dynamic processes such as Berry pseudo-rotation or turnstile rotation (*38*). On the basis of ligand design approaches for high-valent transition metals (*39*), we hypothesized that the lone pair of the S-bound oxygen would become a weak ligand for the Bi center, providing stabilization of putative Bi(V) intermediates. The electron-withdrawing nature of the sulfone group at the ortho position is also expected to render the Bi center more electrophilic, thereby prone to transmetallation and reductive elimination. Additionally, binding two of the three anionic

Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470, Mülheim an der Ruhr, Germany.
*These authors contributed equally to this work.
†Corresponding author. Email: cornella@kofo.mpg.de

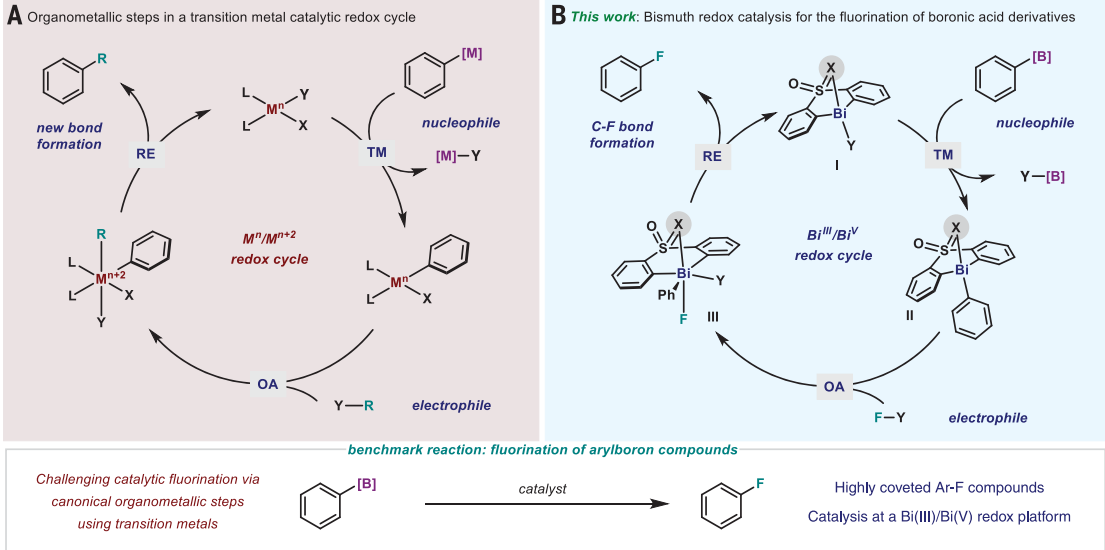


Fig. 1. Can Bi mimic transition-metal behavior in electrophilic fluorination? (**A**) General catalytic two-electron redox cycle of a transition metal. (**B**) Development of an electrophilic fluorination of boronic acid derivatives through a catalytic Bi redox process. L, ligand; M, metal; n, oxidation number; R, organic residue; X, halogen atom (A) or non-anionic ligand (B); Y, anionic ligand.

ligands in a cyclic framework would favor selective reactivity of the unrestrained phenyl group.

With these considerations in mind, we synthesized bismine (**1**) and attempted its oxidation to Bi(V) from the $6s^2$ orbital, through reaction with XeF_2. A smooth conversion to high-valent Bi(V) species was achieved (>95% yield), and x-ray crystallographic analysis revealed a symmetric Bi(V) dimeric structure (**2**) (Fig. 2A). Each Bi atom in **2** adopts a distorted octahedral geometry, with two fluorine atoms positioned trans to each other [F1–Bi1–F2 and F3–Bi2–F4, 164.2(2)°], one of them bridging to the other metal center. The bridging Bi–F bonds are 0.47 Å longer than the terminal bonds. Moreover, the O atoms of the sulfone moiety interact with the Bi centers [Bi1–O1 and Bi2–O2, 3.430(6) Å], thereby forcing the F atom away from linearity and engaging in binding with another Bi, forming a dimer in the solid state. When **2** was heated to 110°C in $CDCl_3$, an unexpected 45% yield of fluorobenzene (**3**) was obtained. This reductive elimination of C–F bond stands in stark contrast to previous reports, in which attempts to thermally induce C–F bond formation from Bi(V) fluoride compounds resulted in decomposition or traces of fluorinated arenes (*40*, *41*). On the basis of the crystallographic information, we anticipated that tuning the electronic properties of the sulfone could influence the Bi(V) center, thus affecting the C–F bond-formation step. Indeed, with a NMe (Me, methyl) group in place of one of the O atoms in the sulfone group (**4**), a notable reduction in yield was observed (**3**, 28%, Fig. 2B). However, when the methyl group in **4** was replaced with CF_3 (**5**), a 94% yield of fluorobenzene (**3**) was obtained after thermal decomposition of the corresponding Bi(V) intermediate. Both **4** and **5** were characterized by x-ray crystallography (Fig. 2B). Despite their comparable Bi–N distances [3.055(2) Å in **4** and 3.038(3) Å in **5**], the electronic nature of the CF_3 group clearly has a notable promotional effect on the reductive elimination process. The elimination of fluorobenzene (**3**) was accompanied by the smooth formation of the corresponding fluorobismine (**6**), which was confirmed by x-ray crystallography. At this point, we presumed that the productive elimination of fluorobenzene from **5** could be ascribed to the low propensity of the NCF_3 group to coordinate to Bi after oxidation, which results in a monomeric Bi(V) difluoride complex. Although attempts to crystallize Bi(V) difluoride compounds derived from **4** and **5** were unsuccessful, the installation of a methyl residue at the ortho position, with respect to the Bi center (**7**), permitted the crystallization of difluorobismine (**8**) (Fig. 2C). X-ray analysis of **8** revealed a monomeric pentavalent complex with trigonal bipyramid geometry (TBP), in which both F atoms occupy apical positions. This geometry is in agreement with the polarity rule for TBP complexes, which predicts that the most electronegative and least sterically demanding substituents are always placed in apical positions—in this case, the fluorides (*42*). As anticipated, coordination of the pending group in **8** is weaker [Bi–N 3.566(4) Å] than in the dimeric sulfone counterpart **2** [Bi–O 3.430(6) Å].

We then turned our attention to the mechanism of this unusual C–F bond reductive elimination. Several para-substituted arylbismine complexes were synthesized (**5** and **9** to **13**, Fig. 3A) and oxidized with XeF_2 to the corresponding pentavalent bismine species (**14** to **19**, Fig. 3A). Thermal decomposition of **14** at 90°C exhibited a first-order kinetic profile (k_{obs} = 1.3 × 10^{-4} s^{-1}, where k_{obs} is the apparent reaction rate constant), in which fluorobenzene (**3**) was produced at the same rate as fluorobismine (**6**) (k_{obs} = 1.0 × 10^{-4} s^{-1} and k_{obs} = 1.1 × 10^{-4} s^{-1}, respectively), reaching 94% yield after 7 hours (Fig. 3A1). An Eyring analysis over a 25°C range revealed a small enthalpy barrier ($\Delta H^{\ddagger}$ = 15.5 ± 0.7 kcal mol^{-1}, where $\Delta H^{\ddagger}$ is the change in enthalpy between reactants and transition state) but a surprisingly high entropic contribution [$\Delta S^{\ddagger}$ = –34.7 ± 1.9 entropy unit (1 e.u. = 1 cal K_{-1} mol_{-1}), where $\Delta S^{\ddagger}$ is the change in entropy between reactants and transition state]. A large and negative value of the entropy parameter is consistent with an associative process in the transition state, although cationic pathways with high degrees of solvent reorganization have also been postulated (*43*, *44*). Hammett kinetic analysis of the thermal decomposition of **14** to **19** to aryl fluorides (**3** and **20** to **24**)

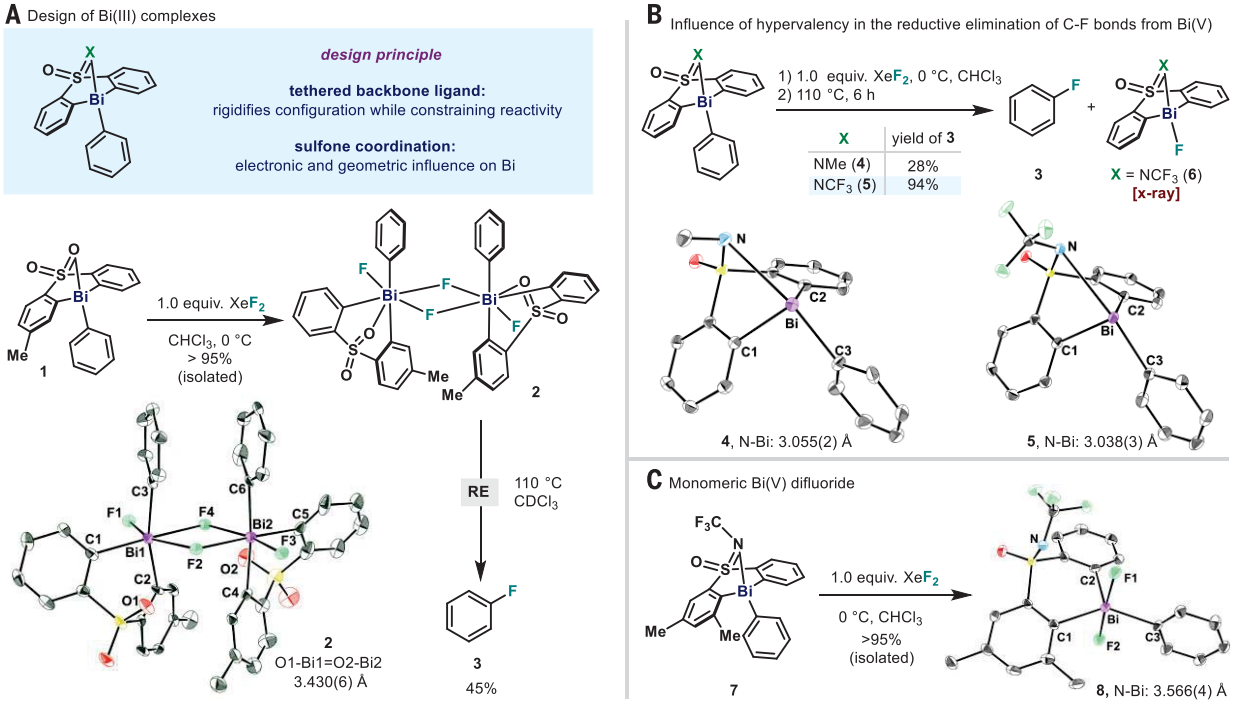


Fig. 2. Design principle and proof of concept. (**A**) Design of the bismine complexes to enable C–F reductive elimination. (**B**) Oxidative addition and reductive elimination sequence to forge C–F bonds. (**C**) Trigonal bipyramidal monomeric Bi(V) difluoride bearing NCF_3. Unless specified, yields were calculated by ^{19}F NMR.

and **6**, revealed a moderately positive slope (ρ = 0.43) when relative constants were plotted versus σ_p^+ (substituent constant), showing good linearity (R^2 = 0.9735, where R_2 is the coefficient of determination and R is the coefficient of correlation). This indicates a slightly faster reductive elimination when electron-withdrawing groups are present in the aryl ring (*39*, *45*). Furthermore, a slower reaction rate for the thermal decay of **14** at 90°C in $CDCl_3$ was obtained in the presence of 1.0 equivalent (equiv.) of Bu_4NF (Bu, butyl group) ($k_{obs} \approx 3.5 \times 10^{-5}$ s^{-1}, partial decomposition observed). Thus, the highly negative entropic value, the excellent linearity obtained when resonance effects are considered, and the slower rate in the presence of fluoride anions suggest a cationic intermediate **14-cation** (Fig. 3B) in equilibrium with **14** (*39*, *43*, *45*). From

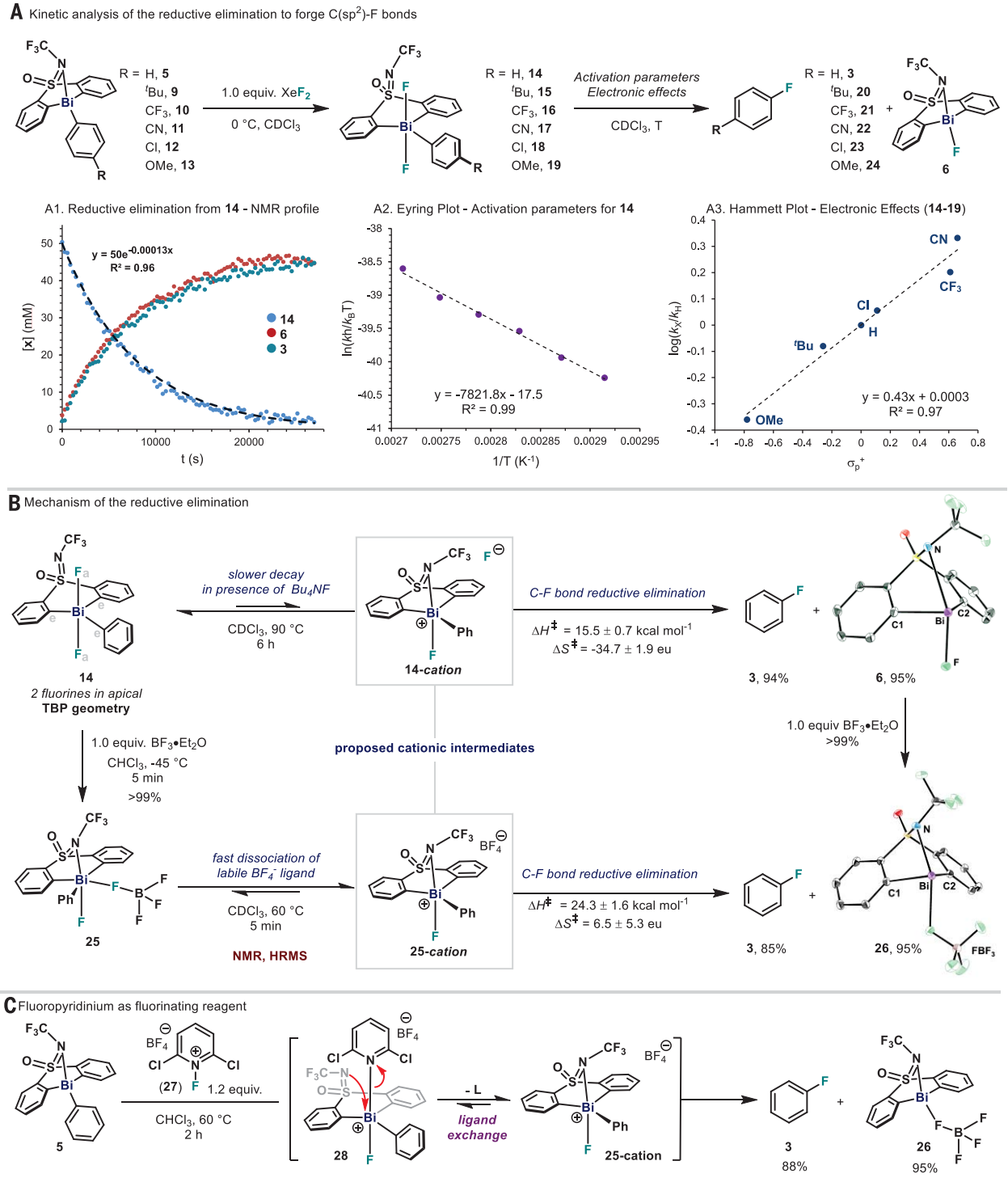


Fig. 3. Mechanistic studies of the reductive elimination. (**A**) Kinetic profile of the reductive elimination from **14**, Eyring analysis, and Hammett studies. (**B**) Experimental evaluation of the reductive elimination step. a, apical position; e, equatorial position. (**C**) Use of fluoropyridinium salts as fluorinating agents. h, hours; k, reaction rate constant; k_B, Boltzmann constant; T, temperature; k_X, reaction rate constant with para-substituent X; k_H, reaction rate constant with para-substituent H; eu, entropy unit.

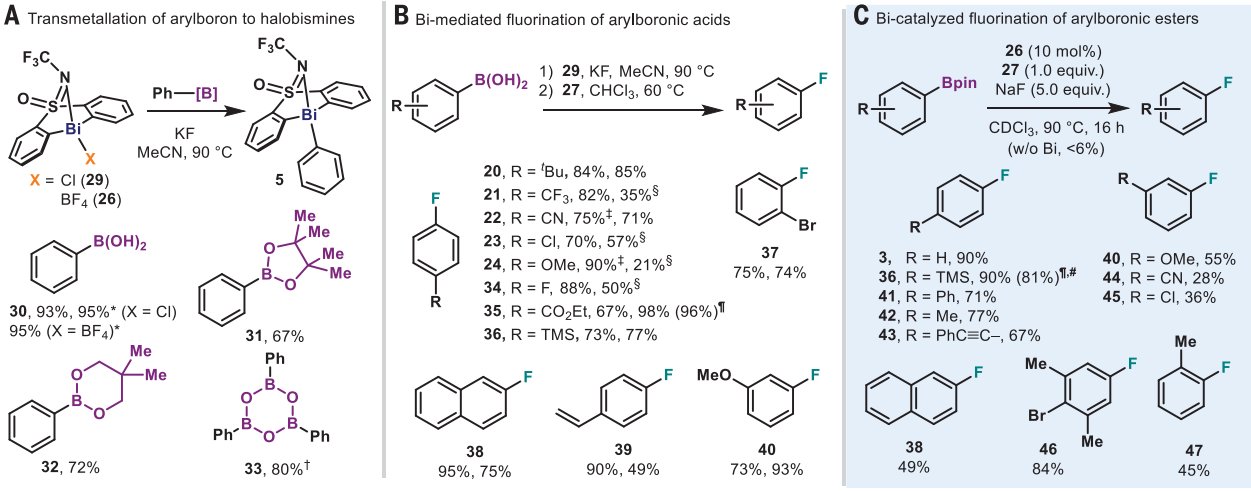


Fig. 4. Bi-mediated fluorination of organoboron compounds. (**A**) Transmetallation of boronic acid derivatives to a Bi(III) complex. Standard conditions: bismine (1.0 equiv.), arylboronic acid derivative (2.0 equiv.), KF (3.0 equiv.), CH_3CN, 90°C, 16 hours. (**B**) Two-step method for the fluorination of boronic acids. Yields are given for step 1 (isolated) and step 2 (determined by ^{19}F NMR), respectively. Step 1: chlorobismine **29** (1.0 equiv.), arylboronic acid (2.0 equiv.), KF (3.0 equiv.), CH_3CN, 90°C, 16 hours. Step 2: arylbismine (1.0 equiv.), **27** (1.1 equiv.), $CHCl_3$, 60°C, 6 hours. (**C**) Catalytic fluorination of aryl boronic esters. Standard conditions: **26** (10 mol%), **27** (1.0 equiv.), arylboronic pinacol ester (3.0 equiv.), NaF (5.0 equiv.), $CDCl_3$, 90°C, 16 hours. Yields determined by ^{19}F NMR. * denotes use of $CHCl_3$ as solvent. † denotes use of 0.66 equiv. of **33**. ‡ denotes use of K_2CO_3 as base. § denotes reaction performed at 110°C. ¶ denotes yield of isolated material after column chromatography. # denotes isolated product contains ~5% of proto-deborylated arene. ^{t}Bu, *tert*-butyl; TMS, trimethylsilyl.

this, elimination of fluorobenzene (**3**) is proposed, with concomitant formation of fluorobismine (**6**). In **14-cation**, coordination of the NCF_3 group to the Bi is expected, providing these Bi intermediates with the adequate balance between stabilization and electronic perturbation to enable C–F bond formation. Reductive elimination from pentavalent complexes with TBP geometry is usually governed by orbital symmetry rules; the Woodward-Hoffmann model (*46*–*48*) predicts that the coupling of equatorial-equatorial and apical-apical ligands should be much more favorable than the equatorial-apical coupling apparently observed in PhF (Ph, phenyl group) elimination from **14**. Recently, Paton and McNally have reported the possibility of bypassing these rules in pentavalent TBP P-based compounds via alternative mechanisms (*49*). Therefore, the formation of a cationic Bi intermediate represents another example which circumvents this constraint (*50*).

To assess the formation of cation intermediates experimentally, we speculated that a Lewis acid could coordinate the apical fluorine syn to the NCF_3 group in **14** and thereby notably elongate the Bi–F bond (*51*), leading to a compound that is geometrically similar to **14-cation**. When **14** was mixed with 1.0 equiv. of $BF_3{\cdot}OEt_2$ (Et, ethyl group) at –45°C (Fig. 3B), exclusive formation of **25** was confirmed in solution by ^1H, ^{11}B, ^{19}F, and ^{13}C nuclear magnetic resonance (NMR) and high-resolution mass spectrometry (HRMS). When this complex was heated to 60°C, fluorobenzene (**3**) was formed in just 5 min, with concomitant formation of tetrafluoroborate bismine (**26**). To further confirm the nature of **26**, 1.0 equiv. of $BF_3{\cdot}OEt_2$ was added to **6**, and immediate conversion to **26** (whose structure was confirmed by x-ray crystallography) was observed. In this case, because of the weakly coordinating nature of the tetrafluoroborate ligand, formation of **3** and **26** occurs readily from the cationic species **25-cation**. An Eyring analysis of the reductive elimination from in situ generated **25** (fig. S11) revealed a high enthalpic contribution to forge **3** and **26** ($\Delta H^{\ddagger}$ = 24.3 ± 1.6 kcal mol^{-1}). In contrast to complex **14**, the reductive elimination from this cationic intermediate showed a minimal entropic contribution ($\Delta S^{\ddagger}$ = 6.48 ± 5.3 e.u.) (*39*). Taken together, these results provide additional evidence for the intermediacy of a cationic species in the reductive elimination from the pentavalent difluorobismine **14**. With this mechanistic information in hand, we hypothesized that a milder and more synthetically useful fluorinating oxidant could afford similar intermediates. Oxidation of Bi(III) compounds to high-valent Bi(V) fluorides has been limited to strong fluorinating agents, such as XeF_2 or F_2 (*40*, *41*). However, when complex **5** was mixed with 1.2 equiv. of 1-fluoro-2,6-dichloropyridinium **27** in $CHCl_3$ (*52*), smooth C–F bond formation occurred at 60°C (Fig. 3C). The high conversion was ascribed to the high lability of the neutral and sterically encumbered 2,6-dichloropyridine ligand after oxidation (**28**), thus enabling coordination of the lone pair from the NCF_3 handle. The resulting intermediate (**25-cation**) can then eliminate fluorobenzene (**3**) and form the corresponding bismine (**26**).

With the goal of turning over the catalytic cycle, we then focused our attention on the transmetallation process between organoboron compounds and chlorobismine (**29**) (Fig. 4A). Whitmire reported the possibility of transmetallating highly nucleophilic tetrarylborates with Bi salicylate salts (*53*). However, transmetallation of less-nucleophilic organoboron compounds to (pseudo)halobismines remains a challenge. Capitalizing on the use of KF as activator, when boronic acid (**30**) was mixed with **29**, smooth transmetallation took place (93% yield). When tetrafluoroborate bismine (**26**) was subjected to transmetallation, a 95% yield of **5** was obtained. Under the same conditions, other organoboron compounds such as PhBpin (pin, pinacol group) (**31**, 67%), PhB(neop) (neop, neopentyl glycolato group) (**32**, 72%), and $(PhBO)_3$ (**33**, 80%) were converted in good yields to phenylbismine (**5**), demonstrating versatility. With the aim of exploring the scope of a two-step method for fluorination, transmetallation of various phenylboronic acids was surveyed. High yields of arylation were obtained (67 to 90%), independently of the functional group in the aryl ring. The resulting arylbismines were then oxidized with **27** and reacted, as above, to release the corresponding arylfluorides (Fig. 4B). The protocol proved general with a variety of *para*-substituted arylfluorides including *tert*-butyl (**20**, 85%), trifluoromethyl (**21**, 35%), cyano (**22**, 71%), chloride (**23**, 57%), methoxy (**24**, 21%), fluoride (**34**, 50%), ester

(**35**, 96%), and trimethylsilyl (**36**, 77%) substituents. Furthermore, *ortho*-bromide (**37**, 74%), naphthalene (**38**, 75%), vinyl (**39**, 49%), and *meta*-methoxy (**40**, 93%) groups could also be accommodated.

We then turned our attention to merging all these steps into a catalytic cycle. On the basis of our initial hypothesis (Fig. 1B), transmetallation of an arylboronic acid derivative to an electrophilic Bi(III) center (**I**) would deliver the aryl bismine (**II**). Subsequently, **II** could be oxidized by the electrophilic fluorine source **27** to furnish a high-valent Bi(V) compound containing both fluoride and tetrafluoroborate ligands (**III**). Rapid decomposition of this intermediate would forge a C–F bond and **I**, thereby restoring the catalyst. Indeed, using tetrafluoroborate bismine (**26**), a catalytic protocol based on Bi for the oxidative fluorination of arylboronic esters was successfully implemented. After a short optimization of the fluoride source, required to activate the boronic ester, a variety of ArBpin were smoothly converted to the corresponding aryl fluorides (Fig. 4C). Using 10 mole % of bismine (**26**), phenylboronic acid pinacol ester (**31**) afforded a 90% yield of fluorobenzene (**3**). Substitution in *para*-position was also tolerated, as exemplified by the presence of trimethylsilyl (**36**, 90%), phenyl (**41**, 71%), methyl (**42**, 77%), alkynyl (**43**, 67%), and bromo (**46**, 84%) groups. Substitution of the aryl group at the *meta*-position presented more difficulties, affording moderate yields of aryl fluoride: methoxy (**40**, 55%), cyano (**44**, 28%), and chloride (**45**, 36%). π-Extended aromatics (**38**, 49%) and sterically hindered substitution, such as a Me group at the ortho position (**47**, 45%), were also amenable for fluorination. The Bi catalyst was essential for the reaction to proceed.

The design presented here enables a Bi complex to undergo transmetallation, oxidative addition with a mild fluorinating agent, and C–F reductive elimination to deliver aryl fluoride compounds. A detailed study of each step paved the way to the development of a catalytic cycle based on the Bi(III)/Bi(V) redox couple, a feat that remained elusive for Bi until now. This mode of reactivity for Bi represents a step forward in mimicking transition-metal–like behavior by an element outside the *d*-block. Although still limited to considerably oxidizing electrophiles and narrow substrate scope, the possibility of performing redox processes with Bi complexes by careful tuning of the ligand properties could potentially be expanded to other similar scenarios, in which a catalyst maneuvers between different oxidation states—a property traditionally associated with transition metals.

REFERENCES AND NOTES

1. A. de Meijere, F. Diederich, Eds., *Metal-Catalyzed Cross-Coupling Reactions* (Wiley, 2004).
2. J. E. Zweig, D. E. Kim, T. R. Newhouse, *Chem. Rev.* **117**, 11680–11752 (2017).
3. B. Su, Z.-C. Cao, Z.-J. Shi, *Acc. Chem. Res.* **48**, 886–896 (2015).
4. R. J. M. Klein Gebbink, M. Moret, Eds., *Non-Noble Metal Catalysis: Molecular Approaches and Reactions* (Wiley, 2019).
5. R. L. Melen, *Science* **363**, 479–484 (2019).
6. P. P. Power, *Nature* **463**, 171–177 (2010).
7. C. Weetman, S. Inoue, *ChemCatChem* **10**, 4213–4228 (2018).
8. G. C. Welch, R. R. San Juan, J. D. Masuda, D. W. Stephan, *Science* **314**, 1124–1126 (2006).
9. L. Wu *et al*., *Nat. Commun.* **10**, 2786 (2019).
10. A. A. Toutov *et al*., *Nature* **518**, 80–84 (2015).
11. H. Bauer *et al*., *Nat. Catal.* **1**, 40–47 (2018).
12. M.-A. Légaré, M.-A. Courtemanche, É. Rochette, F.-G. Fontaine, *Science* **349**, 513–516 (2015).
13. I. G. Molnár, R. Gilmour, *J. Am. Chem. Soc.* **138**, 5004–5007 (2016).
14. N. L. Dunn, M. Ha, A. T. Radosevich, *J. Am. Chem. Soc.* **134**, 11330–11333 (2012).
15. J. C. Sarie, J. Neufeld, C. G. Daniliuc, R. Gilmour, *ACS Catal.* **9**, 7232–7237 (2019).
16. S. M. Banik, J. W. Medley, E. N. Jacobsen, *J. Am. Chem. Soc.* **138**, 5000–5003 (2016).
17. T. V. Nykaza *et al*., *J. Am. Chem. Soc.* **140**, 15200–15205 (2018).
18. R. Mohan, *Nat. Chem.* **2**, 336 (2010).
19. M. B. Kindervater, K. M. Marczenko, U. Werner-Zwanziger, S. S. Chitnis, *Angew. Chem. Int. Ed.* **58**, 7850–7855 (2019).
20. T. Ollevier, Ed., *Bismuth-Mediated Organic Reactions* (Springer, 2012).
21. B. Ritschel, J. Poater, H. Dengel, F. M. Bickelhaupt, C. Lichtenberg, *Angew. Chem. Int. Ed.* **57**, 3825–3829 (2018).
22. J. Ramler *et al*., *Chem. Sci.* **10**, 4169–4176 (2019).
23. F. Wang, O. Planas, J. Cornella, *J. Am. Chem. Soc.* **141**, 4235–4240 (2019).
24. J. Ramler, I. Krummenacher, C. Lichtenberg, *Angew. Chem. Int. Ed.* **58**, 12924–12929 (2019).
25. R. J. Schwamm, M. Lein, M. P. Coles, C. M. Fitchett, *Chem. Commun.* **54**, 916–919 (2018).
26. S. Yamago *et al*., *Angew. Chem. Int. Ed.* **46**, 1304–1306 (2007).
27. G. Pattison, *Org. Biomol. Chem.* **17**, 5651–5660 (2019).
28. P. S. Fier, J. Luo, J. F. Hartwig, *J. Am. Chem. Soc.* **135**, 2552–2559 (2013).
29. M. Tredwell *et al*., *Angew. Chem. Int. Ed.* **53**, 7751–7755 (2014).
30. Y. Ye, M. S. Sanford, *J. Am. Chem. Soc.* **135**, 4648–4651 (2013).
31. T. Furuya, H. M. Kaiser, T. Ritter, *Angew. Chem. Int. Ed.* **47**, 5993–5996 (2008).
32. T. Furuya, T. Ritter, *Org. Lett.* **11**, 2860–2863 (2009).
33. B. H. Rotstein, N. A. Stephenson, N. Vasdev, S. H. Liang, *Nat. Commun.* **5**, 4365 (2014).
34. A. R. Mazzotti, M. G. Campbell, P. Tang, J. M. Murphy, T. Ritter, *J. Am. Chem. Soc.* **135**, 14012–14015 (2013).
35. J.-P. Finet, Ed., *Ligand Coupling Reactions with Heteroatomic Compounds* (Elsevier, 1998).
36. G. G. Briand, N. Burford, *Adv. Inorg. Chem.* **50**, 285–357 (2000).
37. K. Ohkata, S. Takemoto, M. Ohnishi, K.-Y. Akiba, *Tetrahedron Lett.* **30**, 4841–4844 (1989).
38. I. Ugi, D. Marquarding, H. Klusacek, P. Gillespie, F. Ramirez, *Acc. Chem. Res.* **4**, 288–296 (1971).
39. T. Furuya *et al*., *J. Am. Chem. Soc.* **132**, 3793–3807 (2010).
40. T. Ooi, R. Goto, K. Maruoka, *J. Am. Chem. Soc.* **125**, 10494–10495 (2003).
41. S. A. Lermontov, I. M. Rakov, N. S. Zefirov, P. J. Stang, *Tetrahedron Lett.* **37**, 4051–4054 (1996).
42. E. L. Muetterties, W. Mahler, R. Schmutzler, *Inorg. Chem.* **2**, 613–618 (1963).
43. N. P. Mankad, F. D. Toste, *Chem. Sci.* **3**, 72–76 (2012).
44. A. J. Canty, H. Jin, B. W. Skelton, A. H. White, *Inorg. Chem.* **37**, 3975–3981 (1998).
45. J. M. Racowski, A. R. Dick, M. S. Sanford, *J. Am. Chem. Soc.* **131**, 10974–10983 (2009).
46. R. Hoffmann, J. M. Howell, E. L. Muetterties, *J. Am. Chem. Soc.* **94**, 3047–3058 (1972).
47. R. B. Woodward, R. Hoffmann, *Angew. Chem. Int. Ed.* **8**, 781–853 (1969).
48. K.-Y. Akiba, Ed., *Organo Main Group Chemistry* (Wiley, 2011).
49. M. C. Hilton *et al*., *Science* **362**, 799–804 (2018).
50. J. Moc, K. Morokuma, *J. Am. Chem. Soc.* **117**, 11790–11797 (1995).
51. Y. Matano, S. A. Begum, T. Miyamatsu, H. Suzuki, *Organometallics* **17**, 4332–4334 (1998).
52. G. S. Lal, G. P. Pez, R. G. Syvret, *Chem. Rev.* **96**, 1737–1756 (1996).
53. V. Stavila, J. H. Thurston, D. Prieto-Centurión, K. H. Whitmire, *Organometallics* **26**, 6864–6866 (2007).

ACKNOWLEDGMENTS

We thank A. Fürstner and T. Ritter for insightful discussions and generous support and B. Morandi and C. Obradors for insightful suggestions and for proofreading the manuscript. N. Noethling and R. Goddard are acknowledged for crystallographic data analysis. We also thank the analytical department at the MPI-Kohlenforschung for support in the characterization of compounds. **Funding:** Financial support for this work was provided by Max-Planck-Gesellschaft, Max-Planck-Institut für Kohlenforschung, Fonds der Chemischen Industrie (FCI-VCI), Alexander von Humboldt Foundation (to O.P.), and the European Commission (Marie Skłodowska Curie Fellowship, grant no. 833361, to O.P.). **Author contributions:** J.C. and O.P. conceived the idea. All experiments were conducted by O.P. and F.W. NMR experiments and analysis were conducted by O.P. and M.L. The manuscript was written by O.P. and J.C. The project was directed by J.C. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Crystallographic data for structures **2**, **4**, **5**, **6**, **7**, **8**, **26**, and **29** are available free of charge from the Cambridge Crystallographic Data Center under reference numbers 1949430, 1949432, 1949433, 1949435, 1949434, 1949431, 1956313, and 1949436, respectively.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6475/313/suppl/DC1
Materials and Methods
Figs. S1 to S26
Tables S1 to S22
Spectral Data
References (*54*–*69*)

26 August 2019; accepted 17 December 2019
10.1126/science.aaz2258

Microfluidizer

The LV1 Microfluidizer from Analytik is a high-shear homogenizer that brings outstanding processing capabilities to samples as small as 1 mL. It has been designed to achieve operating pressures up to 30,000 psi for samples ranging from 1 mL–6 mL. Using proprietary fixed-geometry interaction chamber technology, the LV1 can process a wide variety of fluids, such as oil-in-water emulsions, solids-in-liquid suspensions, and cells—including the most difficult yeasts and plant cells—in as few as 1-2 passes. What's more, the process is repeatable and is guaranteed to scale up to pilot and/or production volumes.

Analytik
For info: +44-(0)-1954-232776
analytik.co.uk/product/nanoemulsions-lv1-microfluidizer

Black Microplates

Porvair Sciences has expanded its range of high-quality black microplates that minimize sample degradation by exposure to light, even over long storage periods. Available in a choice of 48-, 96-, and 384-well formats, these black plates are precisely manufactured to applicable ANSI/SLAS dimensions, ensuring complete compatibility with almost all readers and automated equipment. Manufactured from polypropylene, they offer excellent heat- and solvent-resistant qualities and are RNAse/DNase free. To prevent evaporation and hydration of dimethyl sulfoxide (DMSO) solutions, and to completely eliminate light from samples, we also offer a selection of Absorb Max light-blocking sealing films that may be applied to the top of the plates. A split backing and edge cutouts make positioning and applying the sealing film quick and easy. The light-blocking and absorbing properties of our black microplates, used in conjunction with Absorb Max sealing films (P/N 229520), provide a perfect combination for protecting light-sensitive assays or samples during storage.

Porvair Sciences
For info: +44-(0)-1978-661144
www.microplates.com/solid-bottom

Ultrasonicator

The ML230 Focused-ultrasonicator can process up to eight samples simultaneously. Developed for superior precision and accuracy, the engineering design enables trouble-free maintenance and a user-friendly experience. The ML230 offers an intuitive touchscreen control with a friendly, convenient software interface. It is easily maintained with the automated water management system. The instrument also features highly tunable acoustic energy for sample-processing volumes of 5 µL–500 µL, and is calibrated to National Institute of Standards and Technology (NIST)-traceable standards, ensuring high-quality, standardized results.

Covaris
For info: 781-932-3959
covaris.com/products/afa-ultrasonication/ml230

Microplate Heater-Cooler

Hamilton unveils the new Hamilton Heater Cooler (HHC) for heating and cooling of ANSI/SLAS footprint microplates. The HHC easily integrates with any of our automated liquid-handling platforms—including the Microlab VANTAGE Liquid-Handling Platform, Microlab STAR, and Microlab NIMBUS—to increase consistency, walkaway time, and overall workflow efficiency in a wide variety of temperature-dependent applications. For those who prefer manual operation, the versatile HHC also functions as a stand-alone device. The HHC offers a broad operating range of 0°C–110°C, while its sturdy construction includes moisture-prohibitive materials, cable insulation, and a built-in drain to collect and manage condensation. Along with Hamilton's endless combination of automation options and on-deck tools, the HHC enables researchers to expand their research boundaries.

Hamilton
For info: 800-648-5950
www.hamiltoncompany.com

Microvolume Spectrophotometer

Implen announces NanoPhotometer N120, an absorbance-based UV/VIS multichannel spectrophotometer. The instrument's extremely compact, stand-alone design (20 cm × 20 cm) allows for rapid, full-scan analysis of up to 12 samples (2 µL) at a time over a wide concentration range (2 ng/µL–8,000 ng/µL double-stranded DNA). Powered by a 2.4-GHz Intel Celeron processor, the N120 significantly reduces measurement time, delivering reliable high-resolution, full-scan data for the accurate determination of sample-purity ratios in as little as 1.7 s per sample, while reducing operational steps for the user by 33%. It can be easily incorporated into workflows by integration with any major laboratory information management system (LIMS) provider for flexibility and added efficiency.

Implen
For info: 818-748-6400
www.implen.com

Sample Storage for Ultralow Temperatures

The Ziath range of CryzoTraq 2D-barcoded cryogenic tubes sets a new standard for cryogenic storage of biological specimens at temperatures as low as –196°C. Worldwide biobanks and biorepositories support many types of life science research, including genomics and personalized medicine. To ensure the integrity and viability of specimens stored in these sample storage facilities over long periods of time, the specimens are placed in sterile storage tubes and immersed in vapor phase liquid nitrogen. Produced using an innovative two-stage injection molding process, the caps on CryzoTraq tubes deliver outstanding leak resistance. To ensure full sample traceability, all tubes have a 2D datamatrix barcode inserted in the tube base. Available in 2-mL and 5-mL volumes, they are certified free from DNA, DNase, RNase, pyrogens, and ATP. Each lot is sterilized with electron-beam radiation. CryzoTraq tubes are manufactured from medical-grade polypropylene and the caps from medical-grade polyethylene.

Ziath
For info: +44-(0)-1223-855021
www.ziath.com

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

BE A PART OF THE CONVERSATION!

Climate Change | Innovation | Sustainability | Artificial Intelligence

Public Engagement | Technology | Diversity | Psychology | Global and Public Health

Chemistry | Natural Hazards and Disasters | Big Data

and more...

AAAS
ANNUAL MEETING
Seattle, WA | Feb. 13–16, 2020

SEE INSIDE FOR DETAILS:

Topical Lectures | Plenary Lectures
Scientific Sessions | Career Workshops
Flash Talks | Milestone Celebrations

Steven Chu
AAAS President
William R. Kenan, Jr.
Professor of Physics
Stanford University

Dear Colleagues:

Scientific and technological research have been invaluable in developing ways of improving life on Earth. The 2020 AAAS Annual Meeting theme, Envisioning Tomorrow's Earth, considers how efforts and advances in science and technology can respond to the new challenges faced by society.

On behalf of the AAAS Board of Directors, I urge you to join us in Seattle from February 13-16, 2020, where this theme will be considered from interdisciplinary scientific perspectives. The AAAS Annual Meeting is the premier event at which you can network with future collaborators from different disciplines. We look forward to seeing you in Seattle.

CONNECT WITH US!

@AAASmeetings
#AAASmtg

/AAAS.Science

aaas.org/meetings

Join us in Seattle, Washington

Connect with scientific and technology professionals, policymakers, funders, educators, students, and members of the media. Together, explore the many ways that research and innovation are reacting to the challenges of the natural and built word:

- Participate in some of the 140+ scientific sessions in 11 disciplinary tracks covering advances in research and policy
- Attend workshops aimed at developing your career
- Take advantage of networking opportunities with new and familiar collaborators

Face-to-face meetings provide unique and unparalleled opportunities for engagement and networking that, as of yet, cannot be re-created virtually. AAAS recognizes the environmental impact of its and other scientific gatherings, and we are constantly examining ways to mitigate this impact. See the growing list of actions at **aaas.org/meetings/ecoefforts**, and send your own suggestions to **meetings@aaas.org**.

Sponsors As of December 2019

AAAS, publisher of *Science*, thanks the sponsors and supporters of the 2020 Annual Meeting.

European Union

Johnson & Johnson INNOVATION

UK Research and Innovation

UNIVERSITY of WASHINGTON
HOST UNIVERSITY

Bristol-Myers Squibb

FRED HUTCH
CURES START HERE®

Arizona State University

UF UNIVERSITY of FLORIDA

NIH National Human Genome Research Institute

Seattle Children's
HOSPITAL · RESEARCH · FOUNDATION

SCIENCE ON TAP AAAS

Topical Lectures

AI Advances and Aspirations

ERIC HORVITZ
Microsoft Research

ASTRIAGraph: A Crowdsourced Knowledge Graph for Space Domain Awareness and Prediction

MORIBA JAH
University of Texas at Austin

Quantum Cloud Computing

ELHAM KASHEFI
The University of Edinburgh

What Place for Solar Geoengineering and Carbon Removal in Climate Strategy?

DAVID KEITH
Harvard University

Baby Brains and our Neuro-Futures

PATRICIA KUHL
University of Washington

Seas the Day: Science Driving Ocean and Climate Solutions

THE HONORABLE JANE LUBCHENCO
Oregon State University

Transgender Children

KRISTINA OLSON
University of Washington

Beyond Complacency: Renewing America's Endless Frontier

DIANE L. SOUVAINE
Tufts University

JOHN P. MCGOVERN AWARD LECTURE IN THE BEHAVIORAL SCIENCES

Variation is the Norm: Darwin's Population and the Science of Emotion

LISA FELDMAN BARRETT
Northeastern University

SARTON MEMORIAL LECTURE IN THE HISTORY AND PHILOSOPHY OF SCIENCE

The Indigenous/Science Project: Collaborative Practice as Witnessing

ALISON WYLIE
Professor, Department of Philosophy, University of British Columbia

Plenary Lectures

AAAS President's Address

STEVEN CHU
Stanford University

PRESENTED BY

Johnson & Johnson INNOVATION

The Second Antibiotic Era

MARYN MCKENNA
Wired

PRESENTED BY

European Union

The Social Life of DNA and the Need for a New Bioethics

ALONDRA NELSON
Social Science Research Council

Updated program information will be posted on **aaas.org/meetings** as it becomes available.

Symposia

Advocating for the Future

ADAPTING TO CLIMATE CHANGE: IMPROVING CONNECTIONS BETWEEN SCIENCE AND SOCIETY
Organized by Amy Snover and Heidi Roop, University of Washington, Seattle, WA

ANIMAL CONSERVATION IN AFRICA: THE INTERSECTION OF SCIENCE AND SOCIETY
Organized by Bryan Ray and Sacha Vignieri, AAAS, Washington, DC

CULTURAL CONNECTIONS: COMMUNICATING SCIENCE TO COMMUNITIES AND CONGRESS
Organized by Bryn Nelson, Seattle, WA

DETECTING, COMBATING, AND IDENTIFYING DIS AND MIS-INFORMATION
Organized by Nadya Bliss, Arizona State University, Tempe, AZ

LEARNING THROUGH CITIZEN SCIENCE: ENHANCING OPPORTUNITIES BY DESIGN
Organized by Lekelia Jenkins, Arizona State University, Tempe, AZ

LOCAL SOLUTIONS TO GLOBAL CHALLENGES: A COMMUNITY-BASED APPROACH TO INNOVATION
Organized by Karel Haegeman, European Commission Joint Research Centre, Sevilla, Spain

PUBLIC TRUST IN SCIENCE: STRENGTH AND SKEPTICISM
Organized by Cary Funk, Pew Research Center, Washington, DC

SAVING SCIENCE JOURNALISM: ACTIONS FOR SCIENCE COMMUNICATION RESEARCHERS
Organized by Sue Ellen McCann, KQED, Inc., San Francisco, CA; Asheley Landrum, Texas Tech University, Lubbock, TX

SCIENCE ADVOCACY: LOBBYING FOR EVIDENCE
Organized by Sandra D. Mitchell, University of Pittsburgh, PA

SCIENCE COMMUNICATION STRATEGIES FOR BUILDING RELATIONSHIPS WITH THE PUBLIC
Organized by Elizabeth A. McCullough, Pacific Science Center, Seattle, WA

SCIENCE IS POLITICAL, NOT PARTISAN: BEST PRACTICES FOR SCIENCE POLICY ADVOCATES
Organized by Holly Mayton, National Science Policy Network, Virginia Beach, VA

THE REPRODUCIBILITY REVOLUTION: IMPACTS ON SCIENCE, JOURNALISM, AND SOCIETY
Organized by Simine Vazire, University of California, Davis, CA

WOMEN'S LEADERSHIP IN THE GLOBAL SCIENCE SYSTEM
Organized by Roseanne Diab, GenderInSITE, Trieste, Italy

Biomedical Futures

AVERTING CATASTROPHIC BIOLOGICAL INCIDENTS IN THE FUTURE
Organized by Kristin Omberg and Karen Taylor, Pacific Northwest National Laboratory, Richland, WA

BIOINFORMATICS AND AI: INNOVATIVE APPROACHES TO RESEARCH AND DATA STORAGE
Organized by Paul Whaley, Lancaster University, United Kingdom

THE BIOLOGICAL ENGINEERING REVOLUTION: STRATEGIES FOR A SUSTAINABLE SCALE-UP
Organized by Philip Shapira, Georgia Institute of Technology, Atlanta, GA; Carrie Cizauskas, Zymergen Inc., Emeryville, CA

BIOLOGICAL NORMALCY AND HUMAN VARIATION: THREE CASE STUDIES
Organized by Andrea Wiley, Indiana University, Bloomington, IN

DYNAMIC RNA MODIFICATIONS: ROLES IN DISEASE AND ENVIRONMENTAL RESPONSE
Organized by Frederick Tyson, National Institute of Environmental Health Sciences, Durham, NC

EQUITY IN GENOMIC MEDICINE: GIVING VOICE TO INDIGENOUS GENETIC VARIATION
Organized by Wyeth Wasserman, University of British Columbia, Vancouver, Canada

GREEN TOXICOLOGY
Organized by Alexandra Maertens and Thomas Hartung, Johns Hopkins University, Baltimore, MD

HUMAN CELL ATLAS: TRANSFORMING BIOLOGY AND HEALTHCARE
Organized by Emily Mobley, Wellcome Sanger Institute, Cambridge, United Kingdom

LIVING DONORS: HOW VOLUNTEERS GIVE TISSUES AND TIME TO ADVANCE SCIENCE
Organized by Kathy Richmond, The Paul G. Allen Frontiers Group, Seattle, WA; Jennifer Pawlosky, Allen Institute, Seattle, WA

MOLECULAR PROBES AND IMAGING: EXPLORING THE INNER UNIVERSE
Organized by Jonathan Bagger and Paul Schaffer, TRIUMF, Vancouver, Canada

ORAL CANCER: USING THE HPV VACCINE FOR PREVENTION
Organized by Jacques E. Nor, University of Michigan, Ann Arbor, MI

PARTICLE THERAPY: PERSPECTIVES FROM PHYSICS, MEDICINE, AND ECONOMICS
Organized by Charles Clark, Joint Quantum Institute, Gaithersburg, MD; Anice Anderson, Private Engineering Consulting, Carmel, IN

PHARMACEUTICAL MANUFACTURING: SUSTAINABLE INNOVATIONS TO PRODUCE ZERO WASTE
Organized by Donna Huryn, University of Pittsburgh, PA; William Beck, University of Illinois, Chicago, IL

SEQUENCING THE GENOMES OF LIFE: PROTECTING BIODIVERSITY AND SUSTAINING SOCIETY
Organized by Emily Mobley, Wellcome Sanger Institute, Cambridge, United Kingdom

SYNTHETIC BIOLOGY: DIGITAL DESIGN OF LIVING SYSTEMS
Organized by Juerg Brunnschweiler and Marianne Lucien, ETH Zürich, Switzerland

SYNTHETIC BIOLOGY AND THE NEXT-GENERATION BIOECONOMY
Organized by Emily Aurand, Engineering Biology Research Consortium, Emeryville, CA

Climate Futures

AI FOR CLIMATE CHANGE ADAPTATION
Organized by So-Min Cheong, University of Kansas, Lawrence, KS

BOUNDARY SPANNING: ENVIRONMENTAL SCIENCE AND DECISION-MAKING
Organized by Kei Koizumi, Shanghai, China

CARBON ECONOMY: CONSIDERATIONS FOR THE UNITED STATES AND PACIFIC RIM
Organized by Naoki Saito, RIKEN, Yokohama, Japan; Steven Collins, University of Washington, Bothell, WA

CHINA'S ROAD TOWARDS ECOLOGICAL CIVILIZATION AND ENVIRONMENTAL SUSTAINABILITY
Organized by Fahu Chen, Chinese Academy of Sciences, Beijing, China

CLIMATE AND MARINE ECOSYSTEMS: EXPLORATION, PREDICTION, AND PROTECTION
Organized by Scott Doney, University of Virginia, Charlottesville, VA

CLIMATE CHANGE AND CLIMATE EXTREMES: PRECIPITATION AND PROJECTIONS
Organized by Tianjun Zhou and Zhongwei Yan, Chinese Academy of Sciences, Beijing, China

CLIMATE CHANGE MITIGATION: THE ROLES OF INDUSTRY AND THE PRIVATE SECTOR
Organized by Murray Hitzman, University College Dublin, Belfield, Ireland

CLIMATE CHANGE SOLUTIONS: CAPTURING GREENHOUSE GAS EMISSIONS IN LAND
Organized by Sara Nichols, University of California, Davis, CA; Lifang Chiang, University of California Office of the President, Oakland, CA

FOSSIL FUEL PRODUCTION AND CLIMATE CHANGE: ALIGNING GOALS AND POLICIES
Organized by Georgia Piggot, Stockholm Environment Institute, Seattle, WA; Gretchen Goldman, Union of Concerned Scientists, Washington, DC

THE FUTURE OF EARTH'S CLIMATE: A WORLD OF EXTREMES
Organized by Patrik Kolar, European Commission Executive Agency for Small and Medium-sized Enterprises, Brussels, Belgium

THE FUTURE OF SUSTAINABILITY SCIENCE IN THE UNITED STATES
Organized by Josh Tewksbury, University of Colorado and Colorado State University, Boulder, CO; Judit Ungvari-Martin, National Science Foundation, Alexandria, VA

THE GLOBAL CARBON CYCLE: IMPLICATIONS FOR LIFE BEYOND THE 21ST CENTURY
Organized by Forrest Hoffman, Oak Ridge National Laboratory, Oak Ridge, TN; Abigail Swann, University of Washington, Seattle, WA

IMPROVED WILDFIRE RESILIENCE: LEVERAGING SCIENCE AND TECHNOLOGY
Organized by Sarah Brady and Teresa Feo, California Council on Science and Technology, Sacramento, CA

OIL SPILLS: ADVANCES IN MITIGATING IMPACT AND CONSEQUENCES
Organized by Ken Halanych, Auburn University, Auburn, AL; Rita R. Colwell, University of Maryland, College Park, MD

SCIENCE DURING CRISIS: BEST PRACTICES, RESEARCH NEEDS, AND POLICY PRIORITIES
Organized by John Randell and Amanda Vernon, American Academy of Arts and Sciences, Cambridge, MA

USING SOLUTIONS FOR RENEWABLE ENERGY, MARINE ENERGY STORAGE, AND WATER USAGE
Organized by Gail Mattson, Brookhaven National Laboratory, Upton, NY

Digital Futures

COMMUNICATION THROUGH TOUCH: FROM BABIES TO SOCIAL ROBOTS
Organized by Andrew Meltzoff, University of Washington, Seattle, WA

THE DIGITAL TRANSFORMATION OF SOCIETIES: ADDRESSING TECHNOLOGICAL DISRUPTION
Organized by Jeffrey Alexander, RTI International, Rockville, MD; Eswaran Subrahmanian, Carnegie Mellon University, Pittsburgh, PA

GERRYMANDERING AND MATHEMATICS: REDISTRICTING THE NATION
Organized by Karen Saxe, American Mathematical Society, Washington, DC

HOW TO COUNTER INTERSECTIONAL BIASES IN SOCIAL MEDIA
Organized by Rochelle Diamond, National Organization of Gay and Lesbian Scientists and Technical Professionals, Pasadena, CA; Jon Pincus, A Change Is Coming, Bellevue, WA

READING EMOTIONS FROM FACIAL EXPRESSIONS: IMPLICATIONS FOR TECHNOLOGY AND HEALTH
Organized by Ralph Adolphs, California Institute of Technology, Pasadena, CA

RESILIENCE IN THE DIGITAL AGE
Organized by Christiane Rousseau, University of Montreal, Canada; Fred Roberts, Rutgers University, Piscataway, NJ

ROBOTIC COMPANIONS AND THE FUTURE OF AI
Organized by Igor Linkov and Benjamin Trump, United States Army Engineer Research and Development Center, Concord, MA

SOCIAL MEDIA AND THE TRANSFORMATION OF SCIENCE ADVICE
Organized by Sarah Foxen and Chris Tyler, University College London, United Kingdom

Engineering the Future

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING: DESIGNING FOR SAFETY AND SECURITY
Organized by Bill Pike and Courtney Corley, Pacific Northwest National Laboratory, Richland, WA

ARTIFICIAL INTELLIGENCE RESEARCH: A COMMUNITY ROADMAP
Organized by Ann Schwartz Drobnis, Computing Community Consortium, Washington, DC

CATALYSTS FOR ENERGY STORAGE: INSPIRED BY NATURE, BUILT BY SCIENTISTS
Organized by Wendy Shaw and Aaron Appel, Pacific Northwest National Laboratory, Richland, WA

CLEAN AVIATION IN TOMORROW'S WORLD
Organized by Michael Kyriakopoulos and Andrea Gentili, European Commission, Brussels, Belgium

CONSTRUCTING AND PERCEIVING BEAUTY
Organized by Daphne Maurer, McMaster University, Hamilton, Canada

DETECTING LIFE AND EXTRATERRESTRIAL TECHNOLOGIES
Organized by Anthony J. Beasley, National Radio Astronomy Observatory, Charlottesville, VA

ENGINEERING BIOLOGY AND THE PROMOTION OF COMMUNITY SECURITY
Organized by Jeffrey Fortman, Engineering Biology Research Consortium, Emeryville, CA

EARTH-SIZED PLANETS ORBITING OTHER STARS: HOW TO FIND AND STUDY THEM
Organized by Heidi Hammel, Association of Universities for Research in Astronomy, Washington, DC

NEW APPROACHES TO FAIRNESS IN AUTOMATED DECISION MAKING
Organized by Sampath Kannan, University of Pennsylvania, Philadelphia, PA; Ann Schwartz Drobnis, Computing Community Consortium, Washington, DC

NEXT GENERATION COMPUTER HARDWARE
Organized by Ann Schwartz Drobnis, Computing Community Consortium, Washington, DC

SOLAR GEOENGINEERING RESEARCH AROUND THE GLOBE
Organized by David Keith and Mariia Belaia, Harvard University, Cambridge, MA

VIEWING TOMORROW'S EARTH FROM SPACE AND SURFACE
Organized by Jack Kaye, National Aeronautics and Space Administration, Washington, DC

Future Earth Systems

ARCTIC INFLUENCES ON SEVERE WINTER WEATHER
Organized by James Overland, National Oceanic and Atmospheric Administration, Seattle, WA

CONVERGENT SCIENCE: COMBATING AFRICA'S WILDLIFE CRISIS
Organized by Mary Ann Ottinger, University of Houston, TX; Meredith Gore, Michigan State University, East Lansing, MI

ECOSYSTEMS' TRUE VALUE: EUROPEAN EFFORTS TO MAP AND ACCOUNT FOR NATURE
Organized by Josefina Enfedaque, European Commission Research and Innovation Directorate-General, Brussels, Belgium; Joachim Maes, European Commission Joint Research Centre, Geel, Belgium

ENVISIONING OCEAN CLIMATE SOLUTIONS FOR THE NEXT GENERATION
Organized by Stephen Posner, University of Vermont, Burlington, VT; Heather Mannix, COMPASS, Silver Spring, MD

THE FUTURE OF WATER AND HUMAN DECISION-MAKING
Organized by Ruby Leung, Pacific Northwest National Laboratory, Richland, WA; Gary Geernaert, Department of Energy, Germantown, MD

GEOSCIENCE LITERACY AND COMMUNITY RESILIENCE
Organized by Cathryn A. Manduca, Carleton College, Northfield, MN

THE GLOBAL WATER CYCLE: UNDERSTANDING TODAY AND TOMORROW
Organized by Michael Dettinger, United States Geological Survey, Carson City, NV

GRASSLANDS AND SAVANNAS: HUMAN IMPERATIVES AND BIODIVERSITY CONSERVATION
Organized by Michael Hill, University of North Dakota, Farrer, Australia

IS THE COAST TOAST? CASCADIA MEGA-EARTHQUAKES, TSUNAMIS, AND POTENTIAL IMPACTS
Organized by Harold Tobin and Alison Duvall, University of Washington, Seattle, WA

MANAGING WATER: NEW TOOLS FOR SUSTAINABLE DEVELOPMENT
Organized by Sera Young, Northwestern University, Evanston, IL

MARINE MAMMAL HEATH ASSESSMENTS: INFORMING SCIENTISTS, POLICY, AND THE PUBLIC
Organized by Stephen A. Raverty, British Columbia Ministry of Agriculture, Abbotsford, Canada; Mike E. Grigg, National Institutes of Health, Bethesda, MD

NATURE CONSERVATION AND COMPUTATIONAL TECHNOLOGIES: EXPANDING THE SCALE
Organized by Daniel Rubenstein, Princeton University, NJ; Tanya Berger-Wolf, University of Illinois, Chicago, IL

NATURE REMADE: ENGINEERING LIFE FROM THE PAST TO FUTURE WORLDS
Organized by Christian Young, Alverno College, Milwaukee, WI; Michael Dietrich, University of Pittsburgh, PA

OCEAN OUTBREAKS ON A CHANGING PLANET
Organized by C. Drew Harvell, Cornell University, Ithaca, NY

THE REEF CRISIS IN EARTH'S FUTURE
Organized by Jere H. Lipps, University of California, Berkeley, CA

GEOSPATIAL INSIGHTS FOR A SUSTAINABLE ENVIRONMENT
Organized by Conrad M. Albrecht and Sharathchandra Pankanti, IBM Research, Yorktown Heights, NY

WILDFIRE SMOKE AND PUBLIC HEALTH: THE SCIENCE AT THE NEXUS
Organized by Ian Gilmour, United States Environmental Protection Agency, Research Triangle Park, NC

Future Health Strategies

AN AGING WORLD: PITFALLS AND PROMISE
Organized by Phyllis Moen, University of Minnesota, Minneapolis, MN

BEETHOVEN AT 250 AND THE SCIENCE OF MUSIC: EMOTION, MEMORY, AND HEALTH
Organized by Daniel Levitin, Minerva Schools, San Francisco, CA

DIET FOR A SICK PLANET: REDUCING EMISSIONS AND THE BURDEN OF CHRONIC DISEASE
Organized by Lydia Zepeda, University of Wisconsin, Madison, Tacoma, WA; Sean B. Cash, Tufts University, Boston, MA

THE DRUG ABUSE CRISIS: KEY FINDINGS FROM THREE LANDMARK STUDIES
Organized by Linda Teplin, Northwestern University, Chicago, IL

HEALTH DISPARITIES: THE INTERSECTION OF PSYCHOLOGY, HISTORY, LAW, AND MEDICINE
Organized by Sophie Trawalter and Dayna Matthew, University of Virginia, Charlottesville, VA

HUMAN EMBRYO RESEARCH REVISITED: SCIENTIFIC, ETHICAL, LEGAL, AND SOCIAL ISSUES
Organized by Kenneth Evans, Rice University, Houston, TX

INFECTIOUS DISEASE FORECASTING: MODELS AND MACHINE LEARNING
Organized by John Drake, University of Georgia, Athens, GA

THE NEUROSCIENCE OF ADDICTION: POLICY CONSIDERATIONS
Organized by Marina Picciotto, Yale University, New Haven, CT

SUPPORTING THE WHOLE STUDENT: MENTAL HEALTH, SUBSTANCE ABUSE, AND WELL-BEING
Organized by Layne Scherer, The National Academies of Sciences, Engineering, and Medicine, Washington, DC

TACKLING SOCIAL RISK FACTORS, HEALTH, AND CARE
Organized by Arlene S. Ash, University of Massachusetts, Worcester, MA

TRACING HOW SCIENTIFIC INFRASTRUCTURE ACCELERATES DISCOVERY
Organized by Elizabeth Lyons, National Science Foundation, Alexandria, VA

TRANSFORMING GLOBAL AND PUBLIC HEALTH THROUGH NATURE
Organized by Usha Varanasi, National Oceanic and Atmospheric Administration, Seattle, WA; Joshua Lawler, University of Washington, Seattle, WA

Future Societal Ethics

50TH ANNIVERSARY OF THE VIETNAM DRAFT: LESSONS FOR THE SCIENCES
Organized by Tim Johnson, Willamette University, Salem, OR

THE CHANGING IDENTITY LANDSCAPE: MULTIRACIAL, INTERSEX, & TRANSGENDER PEOPLE
Organized by Susan Gelman, University of Michigan, Ann Arbor, MI; Kristina Olson, University of Washington, Seattle, WA

DEATH IN THE 21ST CENTURY: WHAT IS LEFT BEHIND
Organized by Robert O'Malley, AAAS, Washington, DC

ETHICAL CONCERNS WITH ADVANCES IN TECHNOLOGY AND GENETICS
Organized by Subrata Saha and Pamela Saha, University of Washington, Seattle, WA

ETHICAL ISSUES IN ARTIFICIAL INTELLIGENCE
Organized by Joseph Halpern, Cornell University, Ithaca, NY

ETHICAL RISKS OF VOICE TECHNOLOGY: A SOCIOLINGUISTIC PERSPECTIVE
Organized by Emily M. Bender, University of Washington, Seattle, WA

ETHNOGRAPHY, OBSERVATION, AND NATURAL HISTORY: TOOLS FOR ETHICAL SCIENCE
Organized by Michelle Bezanson, Santa Clara University, Santa Clara, CA

EVOLUTION OF CRIMINAL JUSTICE: HOW SCIENCE AND EVIDENCE-BASED POLICY INTERACT
Organized by Trisha Chakraborty, Department of Justice, Washington, DC

EVOLVING CONCEPTS OF SCIENTIFIC INTEGRITY AND PRACTICE
Organized by Jonathan Coopersmith, Texas A&M University, College Station, TX

HOW CONGRESS USES SCIENCE AND TECHNOLOGY FOR POLICY: EMERGING RESEARCH
Organized by Karen Akerlof, George Mason University, Fairfax, VA; Chris Tyler, University College London, United Kingdom

IMMIGRATION, CRIME, AND JUSTICE: A DATA-DRIVEN INVESTIGATION
Organized by William Pridemore, State University of New York, Albany, NY

IMPERILED CULTURAL HERITAGE: MITIGATING THREATS, CO-PRODUCING KNOWLEDGE
Organized by Alyne Delaney, Aalborg University, Denmark

IMPLICIT BIAS, EXPLICIT SCIENCE
Organized by Erin Heath, AAAS, Washington, DC

LEARNING FROM PROTACTILE DEAFBLIND COMMUNITIES: TOWARD A MORE TACTILE FUTURE
Organized by Terra Edwards, Saint Louis University, MO; Diane Brentari, University of Chicago, IL

POLITICAL ANIMALS: BEHAVIOR, KNOWLEDGE, REASON, AND TOMORROW'S POLICYMAKING
Organized by David Mair and Marton Hajdu, European Commission Joint Research Center, Brussels, Belgium

PSYCHOLOGICAL SCIENCE: LESSONS FOR THE LAW
Organized by Nora Newcombe, Temple University, Philadelphia, PA

RESPONDING TO CLIMATE CHANGE: SCIENCE, RELIGION, AND CULTURAL PRACTICES
Organized by Curtis L. Baxter and Lilah Sloane-Barrett, AAAS, Washington, DC

USING DATA AND AI TO DISRUPT SEX TRAFFICKING
Organized by Barbara Mack, National Council of Juvenile and Family Court Judges, Seattle, WA

The Future of Food

CROP DIVERSIFICATION: ENSURING A SUSTAINABLE GLOBAL FOOD SUPPLY
Organized by Raul Zornoza, Technical University of Cartagena, Spain; Lise Paresys, National Institute for Agricultural Research, Thiverval-Grignon, France

FOOD OF THE FUTURE: DEVELOPING NEW, SUSTAINABLE, AND HEALTHY SOURCES
Organized by Jens Wilkinson, RIKEN, Saitama, Japan

FOOD SAFETY REGULATIONS: CARCINOGENS AND THE RELEVANCE OF THE DELANEY CLAUSE
Organized by Mansi Krishan, Danone North America, Louisville, CO; Lisa Navarro, Givaudan Flavors Corporation, Cincinnati, OH

HERBICIDE RESISTANCE: CATALYZING TRANSDISCIPLINARY SCIENCE
Organized by David Ervin, Portland State University, OR

IMPROVING FOOD SECURITY POLICIES: INNOVATIVE AGRICULTURE AND MARKET MONITORING
Organized by Felix Rembold, European Commission Joint Research Centre, Ispra, Italy; Inbal Becker Reshef, University of Maryland, College Park, MD

NOURISHING PEOPLE AND PLANET: BUILDING A BETTER, BALANCED FOOD SYSTEM FOR 2050
Organized by Catherine Woteki, Iowa State University, Ames, IA

SUSTAINABLE AGRICULTURE: LAUNCHING STANDARDIZED MICROBIAL ECOSYSTEMS
Organized by Trent Northen, Lawrence Berkeley National Laboratory, Berkeley, CA; Jo Handelsman, University of Wisconsin-Madison, WI

TOMORROW'S TABLE: PLANT GENETICS AND THE FUTURE OF FOOD
Organized by Pam Ronald, University of California, Davis, CA

USING COMPUTING TO SUSTAINABLY FEED A GROWING POPULATION
Organized by Shashi Shekhar, University of Minnesota, Minneapolis, MN; James Hodson, AI for Good Foundation, El Cerrito, CA

Transforming Future Learning

CITIZEN SCIENCE AND BIG DATA: FROM ENGAGEMENT TO ACTION
Organized by Julia K. Parrish, University of Washington, Seattle, WA

DIVERSITY, EQUITY, AND INCLUSION: RECOMMENDATIONS FROM PHYSICS AND ASTRONOMY
Organized by Arlene Modeste-Knowles and Philip W. Hammer, American Institute of Physics, College Park, MD

FINDING THE LOST EINSTEINS
Organized by Michael Feder, AAAS, Washington, DC

INCLUSION IN THE ACADEMY: EXPLORING IDENTITY AND EQUITY IN THE STEM COMMUNITY
Organized by Lina Dahlberg and Robin Kodner, Western Washington University, Bellingham, WA

INCLUSIVITY AND EQUITY IN COURSE-BASED UNDERGRADUATE RESEARCH EXPERIENCES
Organized by Jeffrey Olimpo, University of Texas at El Paso, TX

MENTORSHIP: TOWARD A CULTURE OF EFFECTIVENESS AND INCLUSIVITY
Organized by Maria Lund Dahlberg, National Academy of Sciences, Engineering, and Medicine, Washington, DC

PUBLIC ENGAGEMENT AT UNIVERSITIES: PATHWAYS FOR INSTITUTIONAL SUPPORT
Organized by John Meyer, University of Washington, Seattle, WA; Emily Cloyd, AAAS, Washington, DC

STEM RESEARCH EXPERIENCES FOR HIGH SCHOOL STUDENTS
Organized by Raja GuhaThakurta, University of California, Santa Cruz, CA; Or Graur, Harvard-Smithsonian Center for Astrophysics, Cambridge, MA

STRENGTHENING SUSTAINABILITY PROGRAMS AND CURRICULA IN HIGHER EDUCATION
Organized by Lida Beninson, National Academies of Sciences, Engineering, and Medicine, Washington, DC

TRANSFORMING HIGHER EDUCATION CULTURE: COORDINATING REFORM WITH AGENTS OF CHANGE
Organized by Barbara Natalizio, The Rita Allen Foundation, Princeton, NJ; Erica Kimmerling, American Academy of Arts and Sciences, Cambridge, MA

WOMEN IN STEMM: ADDRESSING UNDERREPRESENTATION
Organized by Alex Helman and Ashley Bear, National Academies of Sciences, Engineering, and Medicine, Washington, DC

Urban Futures

BIOLOGICAL INVASION FORECASTING: UNITING CLIMATE CHANGE, TRANSPORT, AND TRADE
Organized by Erin Grey, Governors State University, University Park, IL; David Lodge, Cornell University, Ithaca, NY

FACTORIES REIMAGINED: MAKING INDUSTRIAL WORK MORE APPEALING
Organized by Erastos Filos, European Commission Research and Innovation Directorate, Brussels, Belgium; Eija Kaasinen, VTT Technical Research Centre of Finland Ltd., Tampere, Finland

NATURAL HISTORY COLLECTIONS AND BIODIVERSITY'S FUTURE
Organized by Keegan Sawyer and Audrey Thevenon, National Academies of Sciences, Engineering, and Medicine, Washington DC

SMART NEW ENERGY VEHICLES AND INTELLIGENT TRANSPORTATION
Organized by Yingjun Qiao and Hongtao Ren, Chinese Academy of Engineering, Beijing, China

SOCIALLY INTEGRATIVE CITIES: PAVING THE WAY TO URBAN SUSTAINABILITY
Organized by Bernhard Müller, Leibniz Institute of Ecological Urban and Regional Development, Dresden, Germany

URBAN PAVING IS GOING PLACES
Organized by Cesare Sangiorgi, University of Bologna, Italy; Ioannis Bitsios, European Commission Research Executive Agency, Brussels, Belgium

URBAN RESILIENCE AND EMERGENCY RESPONSE: THE CLIMATE CHANGE PERSPECTIVE
Organized by Julie Dirwimmer, Fonds de recherche du Québec, Montréal, Canada

URBAN SUSTAINABILITY: SUPPORT FROM ARTIFICIAL INTELLIGENCE AND BIG EARTH DATA
Organized by Daniele Ehrlich, European Commission, Ispra, Italy

Town Halls

Join a forum of experts—both on and off the stage—to discuss data, trends, and strategies facing the scientific community. Topics include:

- Bringing Scientific Evidence to Meet Local Policy Challenges
- The Economics of Climate Change
- Developing Ethical Guidelines for Science Journalism
- Sexual Harassment
- Balancing Science with Concerns About National Security
- And more

SCIENCE ON TAP AAAS Flash Talks

Brief presentations highlighting scientific findings and programs

AI-powered Ultrasound Diagnoses Pneumonia Faster, Better than Experts

Biodiverse Soil Ecosystems: Influencing Climate, Industry, and Life

Building Resilient Communities During Times of High Risk and Uncertainty

The Cancer Paradox: Mutational Secrets Hidden in the Animal Kingdom

Communication, Preparation, and Response: The Effects of Environmental Threats on Health and Behavior

Crop Mutation Application Under Climate Change Mitigates Environmental Pollution

Discipline-based Education Research: Informing a More Effective Undergraduate STEM Experience

The Future of Earth's Ice: A Human Action Story

Gene Editing Goes Global

The Global Disinformation Index: Finding Those Who Corrupt the World's Information

How Decision-Support Technology Pumps Up a Sustainable Groundwater Ecosystem

Iodide: A Primordial Antiperoxidant for Treating Trauma

Let's talk! Bilingualism as One-Health Approach to Understand Neurocognitive and Social Plasticity

Making Medicines Personal: Is It All in Your Genes?

Non-alcoholic Fatty Liver Disease: It's All in Your Head

Population Health Goals for Tomorrow's Earth

Prevalent Plasticizers, Chemotherapeutic Agents, and Disruptions to Embryo Development

Safe Drinking Water for All: Spotting Bacterial Contamination

Stress, Sex, and Inflammation: Metabolic Mediators in the Brain

Science Careers Workshops

Opportunities to gain advice and strategies from experienced STEM professionals

Building A Responsive Network For Tomorrow's Science Communication Needs

Bull's Eye: Developing Specific Aims for a Successful Research Proposal

Careers in State Science, Technology, and Innovation Policy

Citizens, Scientists, and Elections: How Scientists are Engaging in 2020

Communications, Engagement and Advocacy: The Role of Science in Decision-making

Crafting a Narrative for Your Post-academic Career

Envisioning the Global Professional: Uncovering Career Opportunities Worldwide

Envisioning Tomorrow's STEM Workplace: Think Globally, Act Locally

Envisioning Your Career of Tomorrow, Today

European Union Grants: Shake Up Your Research Abroad

Exploring Careers at National Laboratories

Exploring Diverse Avenues to a Career in Science Policy

Friends of the Science Pod: Podcasting, Outreach, and Professional Networking

From the Bench to the Ballot: Scientists in Elected Office

The (Gross) Anatomy of Responding to Peer Review Commentary

How to Make Compelling Outreach Videos When Your Science Seems Dull

How Youth Can Contribute to the Sustainable Development Goals

Impacting the Research Enterprise Through Careers in Research Development

Learning to Manage and Mentor: Skills for Long-term Success in Science

Make 'em Laugh: Science Comedy to Ignite Curiosity and Increase Self-confidence

Navigating Difficult Situations in Public Science Communication

OUT on the Job Search: Finding a Welcoming Environment

Science in the Public Arena: Informing Decision Makers in High-Profile Settings

Shaping STEM Policy without Changing Careers: Local Government Opportunities

Strategies for Securing Philanthropic Funds in Science and Health

Teaching Inclusively with Evidence-Based Strategies

Milestone Celebrations

Centennial of the 19th Amendment

75 Years of *Science—The Endless Frontier* Report

50 Years of Earth Day

Explore additional programming at the AAAS Expo:

Sci-Mic Studio

Where established and emerging scientific podcasts get a conversational perspective on topics raised in the scientific sessions.

TechTangle

Showcases innovative regional and international technology.

Friday: University-level Rovelympics
Saturday: Lifestyle-oriented robots
Sunday: Digital games

Daily Coffee Breaks

Will be available on a first come, first served basis at 9:30 AM and 3:00 PM.

Exhibitors will host discussions, presentations, and demos.

Advance registration rates are available now through **January 24, 2020.**
On-site registration rates will apply thereafter.

	Advance Rates for AAAS Member for members in good standing	Advance Rates for Non-Member for all other attendees	On-site Rates after 1/24/2020 AAAS Member/Non-Member
General Attendee	$310	$440	$380/480
Postdoc	$135	$260	$135/280
K-12 Teacher	$135	$360	$135/380
Retired Professional	$250	$360	$295/380
Student	$65	$95	$75/105
One-Day	$175	$220	$200/240

REPORTERS: The AAAS Annual Meeting Newsroom will be hosted on EurekAlert! at eurekalert.org/aaasnewsroom

For the most up-to-date program information, please visit aaas.org/meetings **aaas.org/meetings**

10 ways that *Science* Careers can help advance your career

1. Register for a free online account on **ScienceCareers.org**.
2. Search thousands of job postings and find your perfect job.
3. Sign up to receive e-mail alerts about job postings that match your criteria.
4. Upload your resume into our database and connect with employers.
5. Watch one of our many webinars on different career topics such as job searching, networking, and more.
6. Download our career booklets, including Career Basics, Careers Beyond the Bench, and Developing Your Skills.
7. Complete an interactive, personalized career plan at "my IDP."
8. Visit our Employer Profiles to learn more about prospective employers.
9. Research graduate program information and find a program right for you.
10. Read relevant career advice articles from our library of thousands.

Visit **ScienceCareers.org** today — all resources are free

SCIENCECAREERS.ORG

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

Funded by the Municipal Government of Shenzhen, Shenzhen Technology University (SZTU) was officially established in 2018 in Pingshan District of Shenzhen. With the strong financial support and favorable policies from Shenzhen government, Shenzhen Technology University (SZTU) aims to be a high degree university of applied sciences and technologies. The total campus area will be 150 hectares.

Mission

To meet the urgent demand from the advanced manufacturing industry locally and nationally, the future SZTU is obligated to produce talent with spirit of craftsmanship such as senior engineers, designers and architects. SZTU aims to establish itself as an open and innovative international university with Chinese characteristic and global outlook.

Education Mode

By learning from the universities of applied sciences of Germany, our university is geared to meet the demands of high-end manufacturing industry with the orientation of employment and entrepreneurship. It is based on enhancing the capability of engineering, practice and innovation. SZTU will explore a new education mode of applied sciences by joining efforts between industries and universities, between production and teaching, between practice and learning.

Talent Cultivation

In accordance with the principle of producing talents in applied sciences with market orientation, SZTU will push the cooperation between industries and universities to a new height; put more focus on the assessment of practical capability and undergo small-sized teaching; promote the integration among teaching, learning and practical training.

Shenzhen Technology University

Seeks Talents Globally: Shenzhen, China

Specialty Setup

SZTU sets up specialties and courses according to the needs of industrial chain and innovative chain. To meet the demand of pillar industry, strategic emerging industry and future industry in Shenzhen, SZTU seeks to hire faculty in a variety of academic fields (capable to teach in English). We welcome faculty applications in **Electronic Science and Technology, Automation, Mechanical Engineering, Internet of Things (IOT), Computer Science and Technology, Light Source and Illumination, New Energies Science and Engineering, Automotive Service Engineering, Automotive Engineering, Transportation Engineering, Logistics Management, Biomedical Engineering, Industrial Design, Environmental Design, International Business, Applied Physics, German, English, Japanese, Teaching Chinese as a Foreign Language. Details are as follows:**

Disciplines and Majors:

Sino-German College of Intelligent Manufacturing

Optical sensing and optical imaging techniques
Optoelectronics device and system
Electronic device design and applications
Ultra-fast laser technology
Laser micro- and nano-fabrication
Control theory and control engineering
Drive and servo control technology
Industrial robot technology and application
Machine vision
Intelligent instruments
Intelligent Manufacturing and Equipment
Precision motion control
Intelligent molding and 3D Printing

College of Big Data and Internet

IOT Operation and Management
IOT Communication
IOT security
IOT technology
IOT application
Artificial Intelligence Application
Software engineering and Theory
Computational Architecture
Mathematics
Basic Computer

College of New Materials and New Energies

Organic optoelectronic materials, devices and systems
Optical communication technology
Inorganic optoelectronic materials, devices and systems
Energy storage materials and devices
Solar energy system and smart micro-grid
High-efficiency solar cell materials and devices

College of Urban Transportation and Logistics

Vehicle Intelligent Networking Technology
New Energy Vehicle Power Drive and Transmission
Car Safety and Accident Analysis"
Automotive Digital Design
Full Lifecycle Automotive Manufacturing and Service
Automotive Agile Detection and Holographic Diagnostic Technology
Automotive Big Data Asset Management
Vehicle Intelligent Terminal and New Ecological Industry
Intelligent Transportation System Engineering
Transportation Planning Management and Big Data Application
Rail Transit Control and Intelligent Maintenance
Rail Transit Facility Design and Operation Management
Smart Logistics and Supply Chain Management

College of Health Science and Environmental Engineering

Medical Instrument/Medical Electronics/Signal Processing etc.
Nano-Bioprobes and Theranostics
Science and Technology for Major Disease Control
Biomedical Imaging
Biological and Medical Sensor Technologies

College of Creative Design

Product design-Vehicle Design
Product design-Cultural and Creative Product Design
Product Design-Electronic Product Design
Jewelry Design
Interior Design
Landscape Design
Public Art

Business School

Supply Chain Management/Logistics Management
Business Management and Controlling
Finance and Corporate Assessment
International Economy/ Statistics
Business Law

College of Engineering Physics

Key technology and applications of high power laser
Laser accelerator physics and applications
Material properties and diagnostics under extreme conditions
Simulation physics and visualization
Precision Instruments and Diagnostics Technology
College Physics Teaching and Experiments Teaching

College of Humanities and Social Sciences (School of Marxism)

German
English
Japanese
Teaching Chinese as a Foreign Language

How to apply

Qualified applicants are encouraged to submit their **Applications Documents** electronically to **recruitment@sztu.edu.cn** with the title as **"Full Name of Applicant + Job Title + Science"**.

Application Documents should include:

A. A comprehensive curriculum vitae with a list of publications, patents and grants.
(Working & Education experience and other kinds of academic experiences should be included)
B. Scanned copies of highest academic degree and related qualification certificates;
C. A statement of future teaching & research plan.

Applications sent inconsistently with our requirements will be considered as invalid.

Contact Us

For more information about our job vacancies, please visit:
https://ohr.sztu.edu.cn/zp/
You can also contact us by telephone or email:
Telephone: +86 755 23256280
Email: **recruitment@sztu.edu.cn**

POSITION LIST

BIT "Teli Forum for International Young Scholars" Beijing 2020

The 5th Teli Forum for International Young Scholars will be held on March 18-19, 2020 by Beijing Institute of Technology (BIT) in Haidian District, Beijing, China. The purpose of the Teli Forum is to promote scientific exchange among researchers, scientists, and engineers in related fields and enhance cooperation.

Teli Forum welcomes international and domestic young scholars.

Agenda

I. The 5th Teli Forum

March 17, 2020 Registration

March 18-19, 2020 Opening Ceremony and Plenary Session, Parallel sessions, Symposiums.

II. "Welcome to BIT" Season (March 1 to June 30, 2020)

For those who cannot take part in the Teli Forum, we sincerely welcome you to sign up and come to our university at any time during the "Welcome to BIT" season. Deans and professors of relevant colleges will talk with you face to face.

Application Conditions

If you, regardless of nationality, can meet any single point of following requirements, apply now!

- hold a doctoral degree obtained from a world-renowned university, have more than 2 years overseas work experience in well-known universities, scientific research institutions or R&D institutions of well-known enterprises abroad; young scholars who are about to obtain a doctoral degree from overseas well-known universities with outstanding academic performance and great development potential.
- Selected in Overseas "High-level Talent Introduction Program" for young scholars, "Chang Jiang Scholars Program" for young scholars, or "Young Top-notch Talent Candidates", supported by the National Natural Science Fund for Excellent Young Scholars, or outstanding young scholars with considerable level.

About the fees

BIT will arrange accommodation for those scholars who have received invitation letters. All expenses including the round-trip expenses will be covered by BIT (Maximum to RMB 20,000 for each overseas scholar, RMB 5,000 for each domestic scholar).

For those who attend the "Welcome to BIT" season, all expenses will be covered by BIT.

About BIT

Beijing Institute of Technology (BIT), one of the key universities in China since the founding of New China and the first batch of universities which has entered the national "211 Project", "985 Project" and the "Top A World-class University". Engineering, material science, chemistry, physics, mathematics, computer science and social sciences ranked in the top 1% of ESI international disciplines, among which engineering ranked in the top 1‰.

BIT now has more than 3,400 staff, including 23 academicians, more than 200 high-level talents and 35 innovative teams; 2 state level cooperative innovation centers, 9 state level laboratories and 6 state level teaching centers.

BIT is building an international university, and has created a global network for exchange and cooperation. BIT will provide an open and inclusive academic environment and a broad stage for all scholars and talents. BIT sincerely welcome foreign and domestic excellent talents and make joint efforts with all talents!

To Apply & Contact us

I. The 5th Teli Forum

Application deadline: Feb.20, 2020

If you are interested, please send your Application Form to the specific email address of a college or institute (http://www.bit.edu.cn/tzgg17/jzyg2/182058.htm) and copy to forum@bit.edu.cn. The application form and email should be named as "Teli Forum+ College to apply+ Name". BIT will issue invitation letters to the selected invitees in time, after qualification examination and academic level evaluation (before March 2, 2020).

II. "Welcome to BIT" Season

You can register at any time. Please send your Application Form to the specific email address of a college or institute (http://www.bit.edu.cn/tzgg17/jzyg2/182058.htm) and copy to forum@bit.edu.cn. The application form and email should be named as "Welcome to BIT + College to apply+ Name".

Contact us:

Human Resources Department

Office of High-level Talents

Jiang Minwei, Xia Yinqiu

+8610-68914243, +8610-68915772

forum@bit.edu.cn, bitrcb@bit.edu.cn

Scan for application

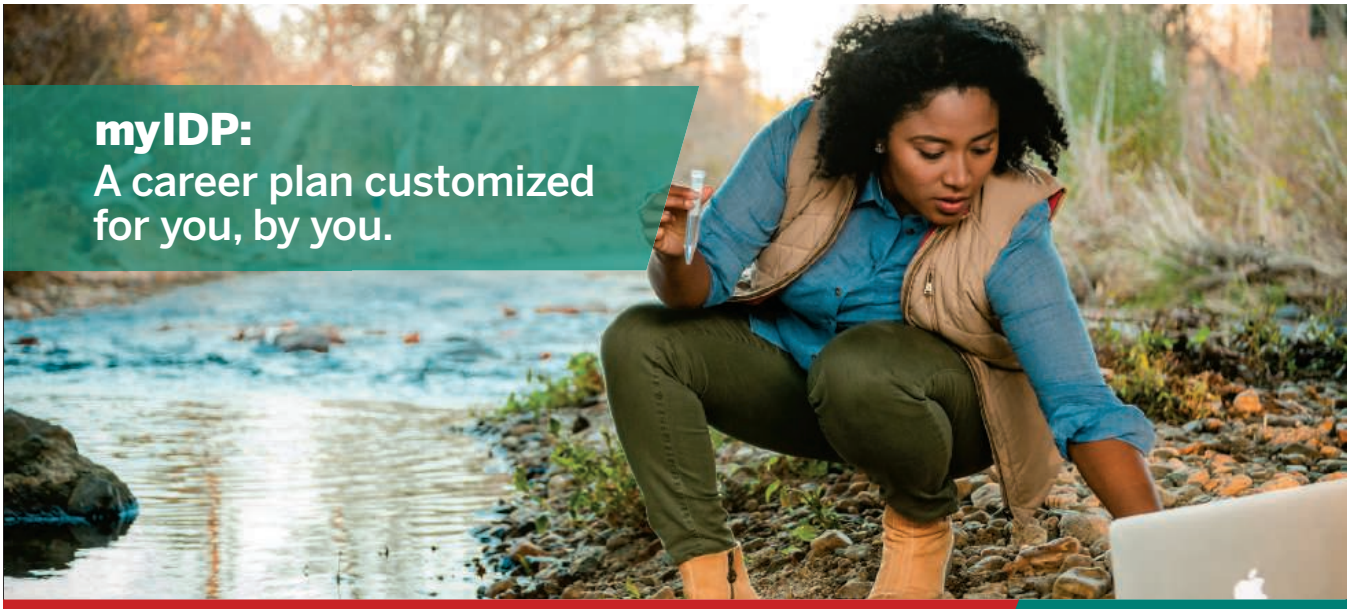
SOMETIMES THE GRASS REALLY IS GREENER SOMEPLACE NEW.

Find your next job at ScienceCareers.org

There's scientific proof that when you're happy with what you do, you're better at what you do. Access career opportunities, see who's hiring and take advantage of our proprietary career-search tools. Get tailored job alerts, post your resume and manage your applications all in one place: sciencecareers.org

ScienceCareers

FROM THE JOURNAL SCIENCE AAAS



For your career in science, 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.
- A tool for setting strategic goals for the coming year, with optional reminders to keep you on track.
- Articles and resources to guide you through the process.
- Options to save materials online and print them for further review and discussion.
- Ability to select which portion of your IDP you wish to share with advisors, mentors, or others.
- A certificate of completion for users that finish myIDP.

myIDP

Visit the website and start planning today!
myIDP.sciencecareers.org

ScienceCareers In partnership with:

FASEB Federation of American Societies for Experimental Biology
UMASS University of Massachusetts Medical School
UCSF University of California San Francisco
UIC
BURROUGHS WELLCOME FUND

ScienceCareers online @sciencecareers.org

By **Erin Gibson**

Getting personal

"Why don't you have a K award?" the interviewer asked, referring to a postdoc award granted by the U.S. National Institutes of Health that is seen by many biomedical scientists as a ticket to a faculty position. It wasn't the first time this question had come up as I interviewed for faculty jobs. But it still left me feeling frustrated. I had published first-author papers in *Science* and *Cell* and acquired plenty of other funding, and I offered the department a unique research plan. I knew I was a strong candidate. Yet in the eyes of some interviewers, my CV was lacking because I did not have an illustrious postdoc award. I struggled through a mental tug of war with myself, wondering whether to shift the conversation and focus on my other accomplishments—or simply tell the truth.

The truth was that when the K award applications were due, I was busy dealing with challenges in my personal life. I waited until after I completed my Ph.D. to become a mother, and I gave birth to our first child during the second year of my postdoc. We were relieved when she was born without any complications, because my first pregnancy, 1 year earlier, had been ectopic, resulting in emergency surgery and the loss of a fallopian tube.

The following year, I became pregnant with our second child. My pregnancy was going well until my 20-week checkup, when the ultrasound technician paused and said, "One moment please. Let me get the doctor." We learned that our baby had abnormal kidney and bladder development and a vascular anomaly in her brain that could be fatal at birth. To our great relief, she survived the cesarean section. But she had urinary problems, her left leg was significantly larger than her right, and she was later diagnosed with a rare genetic disorder.

"The interviews would have been more enjoyable ... had I been able to focus on my science."

The next 2 years of her life were a blur. Taking care of her was a full-time job, requiring hundreds of doctors' appointments and procedures. The procedures eased her health issues somewhat, but the reality is that there's no cure for her condition. It's progressive; it will only get worse.

I am not writing this for sympathy. I am writing this to remind academics that life happens outside the lab. I spent most of my postdoc years struggling to keep up with my research in the midst of major life events at home. Some days, I cried in the corner at work after receiving a worrisome test result from a doctor. Other days, I could barely think because I had stayed up all night with my daughter.

I persisted in academia thanks to the unwavering support of my postdoc mentor, who never forced me to choose between my science and my family. With her encouragement, I put together faculty job applications during my sixth year. I worried that reviewers might balk when they saw I had no prestigious postdoc awards, but I didn't know how to address that. As far as I knew, personal information did not belong in a CV or cover letter.

When this interviewer asked about the K award, however, I decided to be honest. I told him that I didn't have an award because my daughter's health complications had taken priority when the applications were due. He stiffened up and appeared flustered. Clearly that kind of personal disclosure was not the norm. But it didn't seem to hurt me—I ended up receiving an offer from that university and others.

Even so, I wondered whether putting the information in writing as part of my application would have been easier on everyone. In that interview and later ones, I spent an inordinate amount of time talking about my daughter. The interviews would have been more enjoyable and productive had I been able to focus on my science.

I'd like to recommend a solution: When universities advertise job openings, they should invite applicants to describe any events that may have impacted their professional progress. That way, an applicant would feel comfortable explaining that during their 2-year publication gap, they were caring for an ailing parent. Or that they didn't travel to conferences for 4 years because they had two small children at home.

I know that some people will argue that personal information has no place in hiring decisions. But I disagree. To fairly evaluate a scientist's CV, it's important to understand their full journey—children and all. ■

Erin Gibson is an assistant professor at Stanford University.

ILLUSTRATION: ROBERT NEUBECKER

Published by AAAS