Designer immunotherapies for cancer *pp. 990 & 1037*

Mudrocks get rooted *pp. 994 & 1022*

Immune selection for virulence *p. 1030*

Science

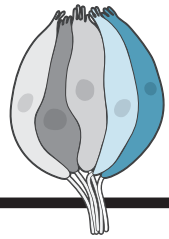
$15
2 MARCH 2018
sciencemag.org

AAAS

FACING UP TO ADVERSITY

How to boost resilience in a challenging world *p. 970*

CONTENTS



991 & 1047
How the tongue tastes sour

2 MARCH 2018 • VOLUME 359 • ISSUE 6379

The roots of resilience

Walling out floodwaters in Bangladesh

FEATURES

NATURE'S STRATEGIES

PERSPECTIVES

SEE ALSO ▸ VIDEO

ON THE COVER

Gopal Mondal stands on the remnants of land at the edge of Polder 32, an artificial island in Bangladesh. Mondal, a resident of Polder 32's Gunari village, lost his land because of river erosion. As natural and human forces threaten homes and livelihoods, resilience—the ability to adapt to adversity, with whatever approach proves most effective—helps people and communities endure. See page 970. *Photo: Tanmoy Bhaduri*

CREDITS: (PHOTO) TANMOY BHADURI; (GRAPHIC) N. DESAI/*SCIENCE*

NEWS

IN BRIEF

IN DEPTH

▸ PODCAST

INSIGHTS

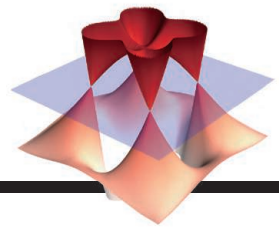
PERSPECTIVES

▸ REPORT P. 1037

▸ REPORT P. 1047

Published by AAAS

CONTENTS



995, 1009, & 1013
Illuminating Weyl points

2 MARCH 2018 • VOLUME 359 • ISSUE 6379

RESEARCH

GRAPHIC: B. YANG *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. Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $1808; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery): $89. 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; 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 is granted by AAAS to libraries and others who use Copyright Clearance Center (CCC) Pay-Per-Use services provided that $35.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Science
AAAS
1200 New York Avenue, NW, Washington, DC 20005
Clarendon House, Clarendon Road, Cambridge, UK CB2 8FH

SUBSCRIPTION SERVICES For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

INSTITUTIONAL SITE LICENSES 202-326-6730 **REPRINTS:** Author Inquiries 800-635-7181 **COMMERCIAL INQUIRIES** 803-359-4578 **PERMISSIONS** 202-326-6765, permissions@aaas.org **AAAS Member Central Support** 866-434-2227 www.aaas.org/membercentral.

Science serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority or conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

INFORMATION FOR AUTHORS See **www.sciencemag.org/authors/science-information-authors**

EDITORIAL

Transparent author credit

Editor-in-Chief, Science *Journals.*
jberg@aaas.org

Authorship on papers is one of the major currencies of the scientific enterprise. Nevertheless, the contributions of different authors to a given paper have remained relatively opaque. Contributions are generally inferred from the order of authors, and implications of position on the authorship list vary between different investigators and scientific fields. A year ago, a group of editors and publishers across a wide range of disciplines met to discuss how to provide a more systemic solution to make author contributions more transparent. This week, their recommendations have been released (www.pnas.org/cgi/doi/10.1073/pnas.1715374115), and I applaud this effort and urge the wide adoption of this system.

More explicit articulation of author contributions can be tremendously beneficial. Earlier in my career, I served on numerous academic promotion committees and, in a few cases, faculty members took it upon themselves to thoughtfully annotate their publication lists, clarifying their roles in individual publications. This allowed the committees to distinguish between those papers for which the faculty member had made major contributions from those involving a lower level of effort. Such distinctions simply were not apparent from the author lists alone. One of the core recommendations involves use of the Contributor Roles Taxonomy (CRediT) system to categorize the efforts of different authors to a paper (docs.casrai.org/CRediT). This system has 14 categories, including Conceptualization, Investigation, Writing (original draft), and Supervision. Each author's efforts can be assigned to one or more of these categories. Many other divisions are certainly possible, but these 14 were derived from practices across a wide range of scientific disciplines. The use of a standardized set of categories can enable human- and machine-readable reporting of these contributions across journals.

Note that CRediT categories are intended to describe author contributions, not to define what constitutes an appropriate contribution to warrant authorship. The group recommended best practices for authorship rules, and we have aligned our editorial policies (www.sciencemag.org/authors/science-journals-editorial-policies) with these recommendations, including the requirement that any author must have made a substantial contribution to the conception or completion of the underlying research, approve the final version of the manuscript, and agree to be accountable for the integrity of his or her contributions and for the overall manuscript. The importance of such rules is to provide a framework for research teams to establish authorship and to avoid the inclusion of inappropriate authors. The committee also makes suggestions for the responsibilities of the corresponding authors of papers. Designation as a corresponding author generally recognizes the individual(s) who played a leading role in the research. As such, this role comes with considerable responsibility when publishing the research, which is articulated in the *Science* family's editorial policies.

"...the new recommendations are a welcome major step."

The group also recommended the use of standard individual identifiers, especially ORCID identifiers (https://orcid.org). A unique identifier associated with each author will enable linking data from different publications, including CRediT contributions, in the future. As someone who has spent considerable effort analyzing large data sets related to the activities of scientists, I can testify to the value of unique identifiers—they decrease the inefficiency and uncertainty associated with sorting out ambiguous names. The overhead associated with collecting such data during the manuscript submission and publication processes can pay big dividends, both to the investigators involved and to the scientific community.

No doubt, a system that promotes authorship transparency will need to evolve as practices are refined and the cultures of different institutions adjust to the approach. But the new recommendations are a welcome major step. The wide adoption of the suggested system should go a long way toward making the roles of different individuals involved with a research paper much clearer, moving beyond the obscure author order code that dominates so much discussion today.

–Jeremy Berg

10.1126/science.aat4136

20% Drop in the rate of gun-related injuries on days when the National Rifle Association held conventions, researchers reported in *The New England Journal of Medicine*. Thousands of gun owners attend the conventions.

IN BRIEF

Edited by **Jeffrey Brainard**

WORKPLACE

U.K. strike may curb research

Some scientists are staying away from their labs and joining picket lines.

In the largest university strike in years, researchers and other academics in the United Kingdom began picketing last week in a disagreement over pensions. One of the biggest pension funds in U.K. higher education is running an estimated deficit of £6 billion, and universities want to solve the problem by shifting from offering defined benefits to making contributions to individual investment accounts. The University and College Union (UCU) in London disputes the projected deficit figure and objects to the change. Strike leaders estimated that tens of thousands of academics will protest over 14 days during the next month at more than 60 institutions, many of them research powerhouses. "It's fair to say there will be a significant impact on research, but how to measure that isn't clear," a UCU spokesperson says. Anecdotal reports suggest many researchers are staying home, and several conferences and seminars have been canceled.

Lawsuit over critique dropped

PEER REVIEW | Environmental engineer Mark Jacobson of Stanford University in Palo Alto, California, last week dropped an unusual, $10 million lawsuit alleging defamation by critics of a journal article he published (*Science*, 10 November 2017, p. 700). Other scientists decried Jacobson's suit, saying such disputes should be handled through scholarly debate, not the courts. In 2015, Jacobson published a paper in the *Proceedings of the National Academy of Sciences* (*PNAS*) arguing that the United States could meet nearly 100% of its energy needs from renewable sources. After *PNAS* published a critique questioning his methodology, Jacobson filed suit in the Superior Court of the District of Columbia against the National Academy of Sciences, which publishes *PNAS*, as well as a co-author of the critique, which Jacobson alleged contained statements of fact damaging to his reputation. In a statement explaining why he dropped the case, Jacobson said: "It became clear … that it is possible there could be no end to this case for years, and both the time and cost would be enormous."

A plan for post-Brexit grants

RESEARCH FUNDING | With negotiations scheduled to begin next month over the United Kingdom's future relationship with the European Union, the London-based Wellcome Trust is calling for the U.K. government to provide access to EU funding sources after Brexit takes effect in 2019. Based on consultations with U.K. and EU research institutions, the report's authors recommend that the United Kingdom pay to be an "associated country" in the Framework Programmes, the main source of competitive grants for the European Union, as Norway and Switzerland do. In return, the United Kingdom should also retain a voice in setting the programs' priorities. The proposal stands a reasonable chance of success in the divorce negotiations, says Peter Tindemans, secretary general of EuroScience, a research advocacy organization in Strasbourg, France.

PHOTO: © GUY SMALLMAN/EYEVINE/REDUX

Drilling for water under the mountains near Cape Town, South Africa, threatens rare plants, ecologists say.

South Africa drilling plan faulted

ECOLOGY | As Cape Town, South Africa, confronts its worst drought in decades, ecologists have criticized a plan to drill into aquifers to supply water, which they say could drive rare plants to extinction within months. Adam West and colleagues at the University of Cape Town argue that tapping groundwater under the Table Mountain group east of the city would lower the water table that sustains dozens of wetland species, including some found nowhere else, such as *Erica bakeri*, a shrublike plant with delicate pink flowers native to a single valley. The mountains are in the Cape Floral Region, a UNESCO World Heritage Site recognized for its biodiversity. Water levels in Cape Town's reservoirs fell so sharply in recent weeks, during the Southern Hemisphere's summer, that officials had warned that the city's water supply would be shut off on "Day Zero" in April, forcing most of the city's 4 million inhabitants to queue for rationed potable water. Although those fears have receded and Day Zero has been put off, officials say tapping the aquifers is necessary to alleviate expected shortages.

NSF to close foreign outposts

INTERNATIONAL AFFAIRS | A plan by the U.S. National Science Foundation (NSF) to close its overseas offices is getting mixed reviews in the scientific community. Last week, NSF announced it will shutter all three of its outposts—in Beijing, Brussels, and Tokyo—by summer. William Chang, who opened NSF's Beijing office in 2006 and now is special adviser for the Asia-Pacific region for the University of Hawaii system, says he's perplexed that NSF would voluntarily sacrifice on-the-ground capacity. "This is really short-sighted," he said. Rebecca Keiser, head of NSF's international office, disagreed and calls the decision "strategic." Shuttering the offices reflects NSF's desire to be nimbler in responding to opportunities "where great science is percolating," she said. That means dispatching small teams on trips of up to a week to explore collaborations. As a pilot test of the new approach, an NSF-led team visited Australia in October 2017 to discuss new research avenues in gravitational wave physics.

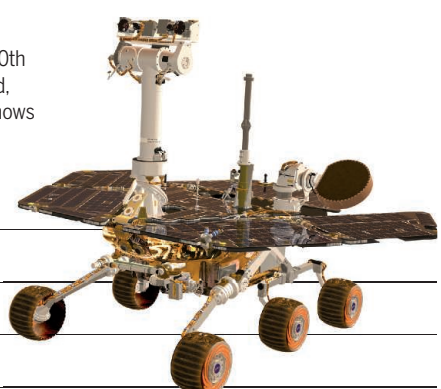
Ethics dispute in Croatia

RESEARCH ETHICS | In an escalating controversy over allegations of scholarly misconduct, one of Croatia's top judges says he has filed criminal complaints against all five members of an ethics panel that found him guilty of plagiarism. In November 2017, the country's Committee on Ethics in Science and Higher Education (CESHE) determined that a 2013 doctoral thesis written by Miroslav Šeparović, president of the Constitutional Court of the Republic of Croatia, contained repeated instances of "incomplete and opaque citations" of other people's work. Šeparović told *Science* he adhered to "unwritten rules" at the University of Zagreb, where he obtained his Ph.D., and that members of the ethics panel are misusing their positions and overstepping their jurisdiction. CESHE Chairperson Ivica Vilibić, a researcher at the Institute of Oceanography and Fisheries in Split, Croatia, said the suit is part of a broader effort to weaken or end the panel, which has embarrassed many politicians and academics since the Croatian Parliament created it in 2005.

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

Surviving on Mars

In February, NASA's Opportunity rover (right) passed its 5000th day investigating the Red Planet. Of NASA probes that landed, only Opportunity and Curiosity still function. This chart shows each probe's landing year and its lifetime in martian days, or sols, which are slightly longer than Earth days.



CREDITS: (IMAGES, TOP TO BOTTOM) HPBFOTOS/ALAMY STOCK PHOTO; NASA; (GRAPHIC) J. BRAINARD/*SCIENCE*; (DATA) NASA

IN DEPTH

ENVIRONMENTAL SCIENCE

Asia's hunger for sand takes toll on ecology

Scientists link species' declines to the mining of construction-grade sand

By **Christina Larson**, *in Kihim, India*

As a morning mist rolls in from the Arabian Sea, young men lead a couple of dozen ox-drawn carts onto a beach south of Mumbai, India's commercial capital. Using shovels and buckets, they pile their rickety wooden transports high with sand, which they will sell to cementmakers. Altering the shoreline is illegal in India, but enforcement of coastal protection zones is lax, says Sumaira Abdulali, a local environmentalist who was beaten up after confronting "sand miners" near here.

Across Asia, rampant extraction of sand for construction is eroding coastlines and scouring waterways. "For a resource we think is infinite, we are beginning to realize that it's not," says Aurora Torres, an ecologist at the German Centre for Integrative Biodiversity Research in Leipzig. "It's a global concern, but especially acute in Asia, where all trends show that urbanization and the region's big construction boom are going to continue for many years." And it is taking an environmental toll that scientists are beginning to assess—and environmentalists hope to reduce.

Already, scientists have linked poorly regulated and often illegal sand removal to declines in seagrasses in Indonesia and in charismatic species such as the Ganges River dolphin and terrapins in India and Malaysia. In eastern China's Poyang Lake, dredging boats are sucking up tens of millions of tons of sand a year, altering the hydrology of the country's largest freshwater lake, a way station for migratory birds. Conservation groups are urging governments to crack down. But the political clout of developers means it will be an uphill—and perilous—battle. Last September, for example, two activists with Mother Nature Cambodia who were filming illegal sand dredging off the Cambodian coast were arrested and convicted of "violation of privacy." They spent several months in jail before being released last month.

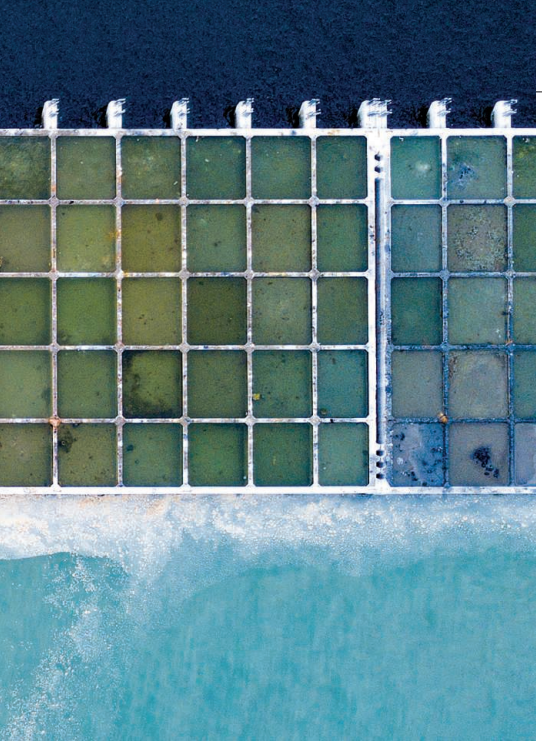
> ***"For a resource we think is infinite, we are beginning to realize that it's not."***
>
> **Aurora Torres**, German Centre for Integrative Biodiversity Research

Used to make concrete and glass, sand is an essential ingredient of nearly every modern highway, airport, dam, windowpane, and solar panel. Although desert sand is plentiful, its wind-tumbled particles are too smooth—and therefore not cohesive enough—for construction material. Instead, builders prize sand from quarries, coastlines, and riverbeds. "The very best sand for construction is river sand; it's the right particle size and shape," says David Shankman, professor emeritus of geography at the University of Alabama in Tuscaloosa, who studies the hydrology of Poyang Lake, a repository of sand deposited by Yangtze River tributaries.

Between 1994 and 2012, global cement production—a proxy for concrete use—tripled, from 1.37 billion to 3.7 billion tons, driven largely by Asian construction, according to a 2014 report from the United Nations Environment Programme (UNEP). Land reclamation projects, too, have a rapacious hunger for sand. Singapore, for example, has expanded its land area by 22% using sand primarily from Malaysia, Cambodia, and Indonesia as fill. All told, UNEP warned, sand mining—on an industrial scale and by individual operators—"greatly exceeds natural renewal rates" and "is increasing exponentially."

Scientists are now tracing the collateral damage. In a paper under review at *Science of the Total Environment*, Richard Unsworth, an ecologist at Swansea University in the United Kingdom, and colleagues explain how sand mining has driven declines of seagrass meadows off of Indonesia. Sediment plumes stirred up by the dredging block sunlight, impeding photosynthesis,

PHOTO: SIM CHI YIN/VII/REDUX



Singapore is attempting to reduce its reliance on imported sand for its land reclamation projects. Much of the fill for a new container port in Tuas, on the island's west coast, is from domestic dredging and excavation.

PHOTO: VERA NIEUWENHUIS ON BEHALF OF PHOTOGRAPHERS WITHOUT BORDERS

his team has found. The meadows nourish several species, including the dugong, which is in decline. "If they lose their food source, the dugong could eventually be goners," Unsworth says.

Another sand mining victim is the southern river terrapin, a critically endangered turtle in Southeast Asia. Every year, Chen Pelf Nyok, a biologist with the Turtle Conservation Society of Malaysia in Kemaman, spends a few weeks patrolling sandy beaches near Malaysia's Kemaman River during the terrapin's brief egg-laying season; her team collects and incubates eggs to protect them from poachers. Three years ago, sand mining erased a nesting site they had monitored. "Terrapin habitat cannot be easily replaced," Chen says, because female turtles return each year to lay eggs at the same beaches.

Sand mining has wiped out nesting sites of critically endangered southern river terrapins in Southeast Asia.

Also under siege, in Bangladesh and India, is the northern river terrapin. "Sand mining is one of the biggest problems and reasons why they are so endangered today," says Peter Praschag, a biologist at the Conservation Breeding and Research Center for Turtles in Graz, Austria. "When the sand banks are gone, the [terrapin] is gone." Other creatures directly affected by river sand mining, scientists say, are the gharial—a rare crocodile found in northern India—and the Ganges River dolphin.

Poyang Lake, a key wintering ground on the East Asian-Australasian Flyway, hosts dozens of migratory species, including almost all of the 4000 or so surviving Siberian cranes. But sand dredging campaigns in the middle Yangtze Basin have expanded rapidly since the early 2000s, when such activities were banned on sections of the lower Yangtze. "Sand mining has significantly lowered the water level, especially in winter," says Lai Xijun, an environmental hydrologist at the Nanjing Institute of Geography and Limnology in China. Falling lake levels can curtail the birds' access to aquatic vegetation. And when lake bottom mud dries and hardens, the birds may not be able to pluck out nutritious tubers.

In grasslands near Poyang, the kind and amount of food the cranes consume "may no longer be enough to fuel egg laying" at the levels the birds managed in the past, says James Burnham, a conservation biologist at the University of Wisconsin in Madison. His group has documented a worrisome decline in the ratio of juvenile cranes to adults at Poyang between 2010 and 2012.

Scientists in China are calling on the government to curtail sand mining across the entirety of the Yangtze Basin. In a letter to *Nature* last October, Yushun Chen of the Institute of Hydrobiology in Wuhan, China, and colleagues argued that sand mining there "has destroyed crucial spawning, feeding and rearing grounds for its aquatic organisms," including the now-extinct Yangtze river dolphin and the endangered Yangtze finless porpoise. "We appeal to the Chinese government to clamp down on this wholesale destruction of aquatic organisms' habitat," they wrote.

"We are not saying we need to stop sand mining altogether. We are saying we need to minimize the impacts," says Jack Liu, a biologist at Michigan State University in East Lansing who is spearheading an effort to assemble a comprehensive picture of the damage. Construction standards should be raised to extend building longevity, he says, and building materials should be recycled. Those sand grains on the beach may not be innumerable after all. ■

Christina Larson is a journalist in Beijing.

ASTRONOMY

Arecibo telescope saved by university consortium

University of Central Florida will take over operations of iconic radio dish

By **Daniel Clery**

After a dozen years of uncertainty about its future, the iconic Arecibo radio telescope in Puerto Rico finally found a savior last week: a consortium led by the University of Central Florida (UCF) in Orlando. The National Science Foundation (NSF) in Alexandria, Virginia, had been looking for another group to shoulder the burden of paying for the Puerto Rican observatory ever since a 2006 review suggested the agency ramp down its funding to free up money for newer projects. "We're delighted that there are signatures on paper," says Richard Green, director of NSF's astronomical sciences division. "That's a fabulous moment at the end of a long process."

Astronomers, planetary scientists, and atmospheric physicists all use the 55-year-old, 305-meter radio dish, the biggest in the world until a 500-meter telescope in China surpassed it in 2016. But its importance has waned. NSF now spends about $8 million a year to run Arecibo, with NASA pitching in an additional $3.6 million. Under the agreement signed last week, NSF's contribution will shrink to $2 million by 2022, with UCF and its partners making up the difference. "There was not a moment's hesitation. It's a real opportunity," says Elizabeth Klonoff, UCF's vice president for research.

UCF will take over management on 1 April, although an agreement detailing the transfer of funds must still be finalized, says James Ulvestad, NSF's chief officer for research facilities. NSF will retain ownership of Arecibo and will regularly review UCF's stewardship of it.

UCF, founded in 1968 to provide technical staff for NASA's burgeoning space program at the nearby Kennedy Space Center, has teamed up with the Metropolitan University in San Juan and Yang Enterprises in

Oviedo, Florida, a company that operates and maintains facilities for NASA and the U.S. Air Force. "Each partner has its own task," says Ray Lugo, head of UCF's Florida Space Institute. He says Metropolitan University will continue its work in education and public outreach, Yang Enterprises will focus on maintenance and operations, and UCF will concentrate on research.

Lugo says the consortium also wants to bring in new customers for the telescope to contribute to costs. He says the Department of Defense may want to use Arecibo to test sensors, while space mining companies could use it to scope out target asteroids. Another possible customer is the privately funded Breakthrough Listen project, which is scanning radio signals from the cosmos for signs of alien intelligence, although Pete Worden, chairman of the Breakthrough Prize Foundation in Menlo Park, California, cautions, "We will need to do a careful analysis of the value and costs."

The consortium also plans to expand the telescope's scientific capabilities, in part by upgrading equipment broken during Hurricane Maria (*Science*, 10 November 2017, p. 704). The agreement with UCF also recognizes Arecibo's significance for Puerto Rico, Ulvestad says. "It's a hugely important technological icon in an underserved community," he says.

Among scientists, relief that the facility avoided closure mingles with regret that NSF is withdrawing its support. "I am pleased by the commitment of new management to continue and to expand the scientific and educational excellence of Arecibo Observatory," says Robert Kerr, a former Arecibo director. But, he adds, "I am disappointed by the tragic and ill-conceived divestment by NSF." Herbert Carlson, a space scientist at Utah State University in Logan who uses Arecibo to study the ionosphere, is optimistic that UCF will make the telescope widely available for science: "I believe it is a good thing for the observatory."

NSF views the agreement with UCF as a possible blueprint for efforts to find alternative funding for other aging telescopes, Green says (*Science*, 11 November 2016, p. 693). In particular, a review committee in 2012 recommended that the agency ramp down its funding for another large radio dish, the 100-meter Green Bank Telescope in West Virginia. "We're hoping that [the Arecibo agreement] will give us and the community confidence that as other divestment efforts proceed, we can reach similar outcomes," Green says. ■

With additional reporting by Adrian Cho.

PHOTO: ISTOCK.COM/SEZERYADIGAR

BIOMEDICINE

Restraining immunity could lower high blood pressure

Researchers hope to launch clinical trial of new strategy

By **Mitch Leslie**

It's fairly easy to give mice hypertension. Just regularly dose them with the hormone angiotensin II. But mixing a molecule called 2-HOBA into the animals' drinking water returns their blood pressure almost to normal, vascular biologist David Harrison of the Vanderbilt University School of Medicine in Nashville and colleagues have found. Now, that observation could open an innovative approach to treating hypertension in people.

Derived from buckwheat, 2-HOBA stands out because of the way it seems to work—by influencing immune cells. "The immune system is an unexpected but important player in hypertension," says vascular biologist Tomasz Guzik of the University of Glasgow in the United Kingdom. Scientists now suspect that immune cells collude with long-recognized culprits such as stress and dietary salt to drive up blood pressure. Safety tests of 2-HOBA in people are already underway, and Harrison, who holds a patent on its use for hypertension, hopes to launch a full clinical trial, which might lead to a new class of treatments that work by restraining the immune system.

More than 1 billion people worldwide have high blood pressure, which promotes heart attacks, strokes, kidney damage, dementia, and other ailments. Current drugs include diuretics that reduce the amount of water in the body and β blockers that decrease how much blood the heart pumps. Yet about 15% to 20% of patients don't improve. "Clearly, we are not managing the condition appropriately at the moment," says vascular biologist Grant Drummond of La Trobe University in Melbourne, Australia.

Scientists first suggested that the immune system modifies blood pressure more than 50 years ago. But a 2007 study by Harrison, Guzik, and colleagues was a watershed. The researchers infused angiotensin II into mice genetically altered to lack two types of immune cells: B cells and T cells. The animals' blood pressure remained about 20 points below that of controls, which also received the hormone. When the researchers restored T cells to the modified rodents, however, their blood pressure surged. That result "was a major finding that triggered the explosion of interest in the field," says nephrologist Thomas Coffman of the Duke-National University of Singapore Medical School.

In 2011, cardiovascular biologist Ernesto Schiffrin of McGill University in Montreal, Canada, and colleagues took the opposite tack. They infused immune-suppressing regulatory T cells into hypertensive mice and reported that the cells reined in blood pressure and reduced the amount of blood vessel damage the animals suffered. "The data, at least in animal models, are compelling that the immune system is involved in hypertension," Coffman says. Some human studies also indicate a connection. In a 2006 paper, for example, researchers revealed that blood pressure declined by more than 10% after patients with psoriasis or rheumatoid arthritis began taking an immune-inhibiting drug.

Researchers doubt that immune cells instigate hypertension. "The immune system probably kicks in after some of the

Activated immune cells may contribute to high blood pressure and offer a target for treatment.

more traditional risk factors are present," Drummond says. These factors, which can include a high-salt diet, stress, and a naturally overactive sympathetic branch of the nervous system, spur an initial increase in blood pressure that damages blood vessels. Immune cells detect that damage, and their response sparks "a vicious circle that leads to the progressive elevation of blood pressure," Schiffrin says.

Among other effects, immune cells disrupt the function of the endothelial layer, the lining of the blood vessels, counteracting "all the good things that the endothelial cells produce," says physiologist Brett Mitchell of Texas A&M College of Medicine in College Station. For example, those cells normally emit nitric oxide, which relaxes blood vessels and reduces blood pressure—and immune cells inhibit nitric oxide production. The cells also wreak havoc in the kidneys, stimulating the organs to hold on to more sodium, which in turn spurs the body to retain more water.

"The question of the decade," Harrison says, has been what switches on the immune cells. His team thinks it has isolated one signal: oxidized lipids known as isoketals that form inside blood cells. In 2014, he and his colleagues discovered that these molecules are unusually abundant in certain immune cells of mice with high blood pressure—and that the same is true in patients with hypertension. Isoketals adhere to and damage proteins, and Harrison's group found that the resulting injured proteins stimulate immune cells known as dendritic cells, which in turn activate T cells. It's "a pretty good case," says nephrologist Richard Johnson of the University of Colorado Anschutz Medical Campus in Aurora.

Harrison's potential blood pressure treatment, 2-HOBA, thwarts isoketals by muzzling their reactive ends. That probably won't impair our defenses against pathogens. But researchers are divided over whether to test the more powerful immune-suppressing drugs that patients take for illnesses such as psoriasis, Crohn disease, and rheumatoid arthritis. Schiffrin argues that these drugs are too risky to use in hypertension, which people can live with for decades. "We don't want to produce fatal symptoms in a patient … because we were playing around with their immune system."

Drummond, however, says such drugs could serve as short-term treatments for people who don't respond to other therapies. "It is so important that we get blood pressure under control," he says, that "there is strong justification" for a clinical trial to test some of these drugs. ■

IMAGE: KC ROEYER/UNIVERSITY OF CALIFORNIA, BERKELEY

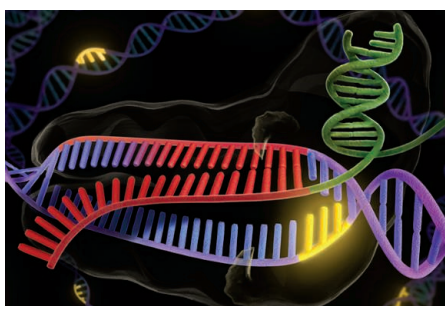
BIOLOGY

Genome editor gets more versatile and precise

Modifying enzyme used by CRISPR brings four times more DNA within reach of its molecular scissors

By **Jon Cohen**

You wouldn't know it from the excitement generated by the revolutionary genome editing method known as CRISPR, but as practiced now, it is far from perfect. Its standard components can find and cut DNA in only a limited fraction of the genome, and its molecular scissors are wobbly, leading to "off-target" mutations. Many groups are trying to do better, and now, a team led by chemist David Liu at Harvard University has engineered a version of CRISPR that potentially is both more dexterous and more precise.



The genome editor CRISPR cuts DNA with help from a guide RNA (green and red) and a Cas9 enzyme (outline) that latches onto a three-base sequence (yellow).

"This is very impressive and important work," says CRISPR pioneer Erik Sontheimer of the University of Massachusetts Medical School in Worcester.

CRISPR comes in many flavors, but they all depend on a guide molecule composed of RNA to carry a DNA-cutting enzyme—the most commonly used one is known by the shorthand Cas9—to a specific stretch of the genome. This complex, however, homes in on DNA landing pads that have specific molecular features. The enzyme in the standard CRISPR toolkit, called spCas9 for its natural source, the bacterium *Streptococcus pyogenes*, can only land on genome segments that have at one end a specific three-base trio: N, where N is any of DNA's four bases, followed by two guanines (Gs). Only about one-sixteenth of the 3.2-billion-base human genome has the right sequence. "That's been a real limitation," Liu says.

The new work, reported online in the 28 February issue of *Nature*, modifies the Cas9 enzyme, creating at least four times as many potential docking sites. In theory, this could allow researchers to, say, cripple or replace many parts of genes associated with human disease that CRISPR currently cannot touch.

Liu's lab began by engineering a large variety of slightly altered spCas9s. The group then selected for ones that could use a broader range of the 64 possible, three-base landing pads—technically referred to as protospacer adjacent motifs, or PAMs. They've dubbed their new enzymes xCas9s, and the best one works with NGN, a sequence that occurs in one-fourth of the genome.

Liu expected that in return for gaining the ability to latch onto more places, xCas9 would pay a penalty: more of the potentially dangerous off-target cuts that concern researchers hoping to unleash CRISPR in medicine. After all, conventional thinking holds that Cas9, naturally part of a bacterial immune strategy, evolved to be as promiscuous in its DNA binding as it could be without compromising specificity. "PAM binding is supposed to be the gatekeeper, and if your gatekeeper is drunk and lets lots of Cas9 into the dance, that should screw up the targeting," Liu explains. But the opposite happened. "If you ask me for a detailed mechanistic explanation for why that is, my answer is, 'I don't know,'" he says.

Stanley Qi, a CRISPR researcher at Stanford University in Palo Alto, California, says this win-win situation is "amazing," and should excite many labs. "The real test here is if people rush to use this xCas9 while forgetting about the original version," Qi says. "At least in my lab, we are very eager to try this out for our applications."

Liu cautions that the standard Cas9 has proved itself over the years; his lab has only tested the new xCas9 on a few dozen sites in the genome so far, compared with the thousands the original has been shown to hit. "I'm not 100% sure xCas9 is going to be flat out better than spCas9," Liu says. "I want everyone to test it because I want to know the answer." ■

METROLOGY

Better atomic clocks herald new era of timekeeping

Metrologists move to redefine the second with a visible light standard that would boost accuracy by a factor of 100

By **Edwin Cartlidge**

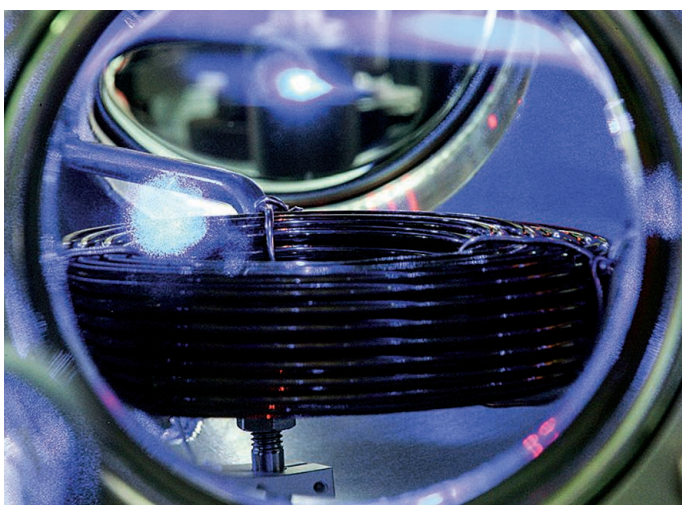
The atomic clocks that mark official time lose the equivalent of just 1 second every 200 million years. But metrologists are not satisfied. A more precise time standard might improve the navigation of spacecraft and help experimenters look for variations in fundamental constants that would signal new physics. So the push is on to replace current clocks, which are tuned to a specific microwave frequency, with even better clocks that exploit higher-frequency visible light.

In a paper out last month, a group of experts set up by the International Bureau of Weights and Measures (BIPM) in Sèvres, France, lays out a road map for the steps needed to redefine the unit of time—the metric second—in terms of optical radiation. Already, physicists at the National Institute of Standards and Technology (NIST) Boulder Laboratories in Colorado appear to have satisfied one of the road map's key requirements—a 100-fold improvement in accuracy over the best microwave clocks—using a pair of optical clocks.

Clocks mark time by tracking a periodic action. A grandfather clock relies on the regular swings of a pendulum, and the original definition of the second was based on the length of a day as fixed by Earth's spin. Current atomic clocks depend on the oscillations of a microwave beam at the precise frequency needed to excite atoms of cesium-133 to a higher energy level. In 1967, the second was defined as 9,192,631,770 cycles of a beam tuned to the cesium standard. Today, the best cesium clocks have accuracies of 1.6 parts in 10^{16}.

The frequency of visible light is about 100,000 times higher than that of microwaves, promising even more precision. However, the lasers needed to cool atoms and provide a stable reference at these frequencies are a challenge to build. Improvements in laser technology led BIPM a few years ago to begin reviewing the accuracy of optical clocks. On 14 February it published a paper in its journal, *Metrologia*, setting out five milestones that should be met before the second can be redefined based on visible light. And in unpublished work, Andrew Ludlow and colleagues at NIST appear to have reached the accuracy stipulated by BIPM's first milestone.

The NIST team operated two optical clocks, using several lasers to cool and trap a few thousand ytterbium atoms in an "optical lattice" and then excite a particular energy transition in those atoms. The researchers found that the two clocks ticked at the same rate to within 1.4 parts in 10^{18}—just over 100 times better than the top cesium devices. "It would be the first time that two clocks of the same species have been shown to agree at that level," Ludlow says.



A cloud of cold strontium atoms, glowing with a blue light, is trapped in the vacuum chamber of an optical clock at Germany's National Metrology Institute.

PHOTO: PHYSIKALISCH-TECHNISCHE BUNDESANSTALT/CC-BY

Being a cautious bunch, metrologists will not accept the NIST result at face value. Jérôme Lodewyck, a physicist working on strontium lattice clocks at the Paris Observatory's Time-Space Reference Systems lab, says the NIST result "is probably correct," but that when it comes to changing the second, "probably correct is not good enough." He notes that something unexpected—perhaps a stray electric field—could throw off two clocks in the same lab by the same amount, making the drift undetectable.

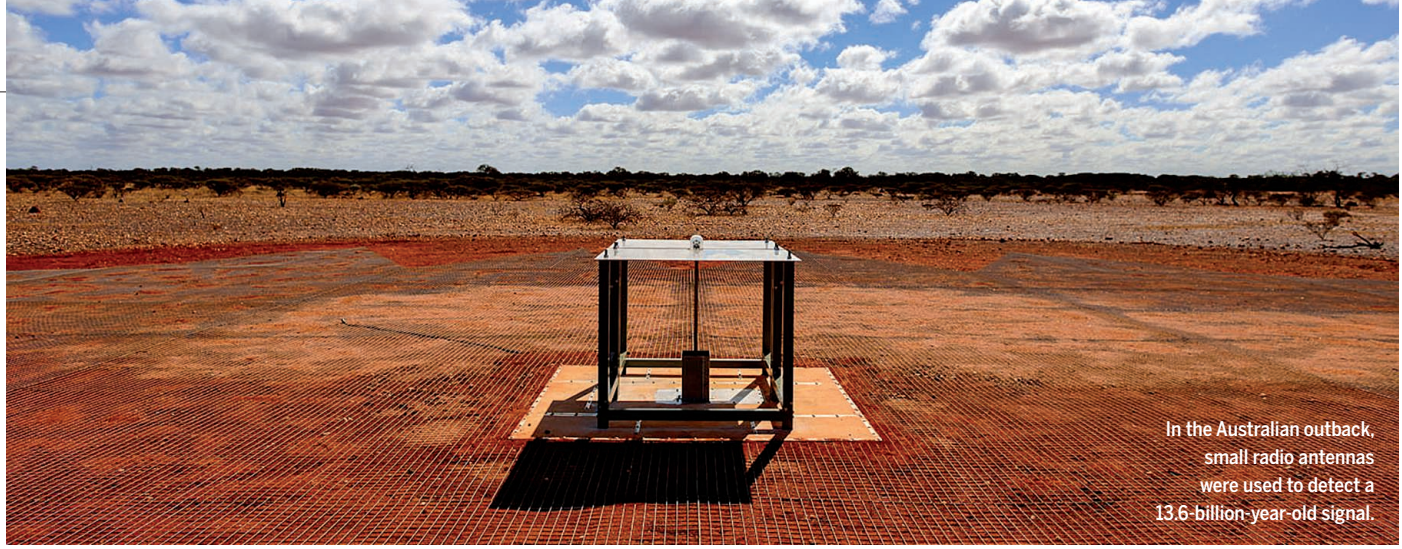
The BIPM road map calls for a number of cross-checks, including reaching the required accuracy with clocks in different labs. Another check involves comparing the ticking of different types of atomic clocks. The NIST scientists and colleagues at JILA, a research institute down the road in Boulder, are doing just that, comparing the ytterbium clocks to others that rely on strontium atoms and aluminum ions. Ludlow says the measurements are not far off the desired accuracy, and that once they've finished they can compare their results with labs in Europe or Asia.

One issue the road map does not address, Ludlow says, is how to choose which atomic transition—and hence type of clock—will ultimately come to define the second. He prefers clocks made from lattices of neutral atoms because they need only run for a few hours before reaching their stated accuracy, whereas rival devices made from single trapped ions might need weeks to do that. But he acknowledges that the jury is still out. "Ion clocks have made great advances in the last decade and they keep getting better," he says.

Patrick Gill, a laser physicist at the National Physical Laboratory in Teddington, U.K., says the switchover to optical clocks shouldn't happen while they are all improving so quickly. He says that officials might agree on a new definition when the world's top metrological body—the General Conference on Weights and Measures—meets in 2026. (The organization meets every 4 years and this year is expected to approve new definitions for four other metric base units: the kilogram, the ampere, the kelvin, and the mole.)

Ekkehard Peik, head of the time and frequency group at Germany's National Metrology Institute in Braunschweig, is more cautious still. Arguing that cesium clocks are accurate enough for today's applications, he says a redefinition can probably wait until the early 2030s, offering more time for scientists to duke it out with competing clocks. Such rivalry, he says, "is what also drives progress and one should not be afraid of it." ■

Edwin Cartlidge is a journalist in Rome.



In the Australian outback, small radio antennas were used to detect a 13.6-billion-year-old signal.

PHOTO: COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION

COSMOLOGY

Cosmic dawn signal holds clue to dark matter

Microwaves from the big bang probe primordial gas clouds as the first stars turn on

By **Adrian Cho**

Using radio antennas the size of coffee tables, a small team of astronomers has glimpsed the cosmic dawn, the moment billions of years ago when the universe's first stars began to shine. The observation also serves up surprising evidence that particles of dark matter—the unseen stuff that makes up most of the universe's matter—may be much lighter than physicists thought.

If it holds up, the result could sharpen cosmologists' picture of the early universe and shake up the search for dark matter. "It's going to generate a huge amount of interest," says Kevork Abazajian, a theoretical cosmologist at the University of California (UC), Irvine. But others worry that the subtle radio signal reported by the team could be an artifact. "I don't think that right now, at least in my mind, it's a clear discovery," says Aaron Parsons, an experimental cosmologist at UC Berkeley.

The data come from the Experiment to Detect the Global Epoch of Reionization Signature (EDGES), a $2 million array of three radio antennas in the outback of Western Australia. The five EDGES researchers searched for signs that the hydrogen atoms that pervaded the newborn universe had absorbed microwaves lingering from the big bang.

The absorption marks the moment just after the first stars began to shine. Before that moment, the atoms' internal states were in equilibrium with the microwaves, emitting as much radiation as they absorbed. But light from the first stars jostled the atoms' innards, disrupting the equilibrium and enabling the atoms to absorb more of the microwaves than they emit.

The expansion of the universe stretches the absorption signal from its original 21-centimeter wavelength to longer radio wavelengths. However, radio noise from our galaxy is 30,000 times more intense. To subtract it, EDGES researchers relied on the noise's smooth, precisely predictable spectrum. This week in *Nature*, they report detecting the tiny absorption signal—the cumulative shadows, they conclude, of hydrogen clouds that existed between 180 million and 250 million years after the big bang.

It's the first thing scientists have seen in the time between the cosmic microwave background, 380,000 years after the big bang, and the oldest known galaxy, which shone 400 million years later, says EDGES leader Judd Bowman. "This is really the only possible probe that we have of the time before the stars," says Bowman, who is an experimental astrophysicist at Arizona State University in Tempe. Ultimately, scientists hope to use the absorption signal or the fainter emission of 21-centimeter radiation from gas clouds at slightly later times to map the 3D distribution of hydrogen during these so-called cosmic dark ages, tracing its evolution into embryonic galaxies.

The absorption is more than twice as strong as predicted, which suggests that the hydrogen was significantly colder than previously thought. The gas must have lost heat to something even colder, and the only colder thing around was dark matter, which was coalescing into the clumps that would seed the formation of galaxies, reasons Rennan Barkana, an astrophysicist at Tel Aviv University in Israel. In a second paper in *Nature*, Barkana argues that to cool the hydrogen, the dark matter particles must have been less than five times as massive as a hydrogen atom. Otherwise the atoms would have bounced off them without losing energy and getting colder, just as a Ping-Pong ball will bounce off a bowling ball without slowing down.

Many dark matter searches have targeted hypothetical weakly interacting massive particles, which are generally expected to weigh hundreds of times as much as a hydrogen atom. As those searches have come up empty, some physicists have begun searching for lighter dark matter particles (*Science*, 24 March 2017, p. 1251). The new result may encourage them, Abazajian says.

However, it's too early to rule out a more mundane explanation for the unexpectedly strong absorption, cautions Katherine Freese, an astrophysicist at the University of Michigan in Ann Arbor. "Is [this scenario] the only way to explain this? Of course not."

A more pressing question is whether the signal is an experimental artifact, Parsons says. The measurements rely on calibrations that could produce false signals if they are off by just a few hundredths of a percent, he says. Bowman says he and his colleagues "have gone as far as we can go to ensure that there isn't an error, but, of course, we're eager for others to confirm the result."

Confirmation could come from other experiments that are probing the dark ages. Parsons leads one, called the Hydrogen Epoch of Reionization Array in South Africa, which is trying not just to detect the faint signals, but to map them across the sky. They may soon show whether cosmic dawn has really broken. ■

THE ROOTS OF RESILIENCE

By **Jennifer Couzin-Frankel**

Resilience is on many people's minds these days. Hurricanes and fires regularly wallop communities. The risks of climate change loom large, and the horrors of war and the refugee crises it spawns show no signs of abating. Bitter political divisions have yielded to acrimony and gloom.

It's an unsettling time—made more so because we humans have nurtured many of these crises, yet feel unable to control them. But even though hardship cripples some, others rebound. What can science teach us about how we might gird for future challenges and adapt to them?

To find out, we followed the stories. They took us to New Orleans, Louisiana, where social scientists are tracking Hurricane Katrina survivors. They took us to the Middle East, where scientists are testing a program designed to foster resilience among traumatized children. They took us to Bangladesh, where residents are rethinking

PHOTO: MARIO TAMA/GETTY IMAGES

At a 2006 vigil for those who died in Hurricane Katrina, mourners hold hands.

whether resilience means bending to nature's will or fighting it. And they took us deep into the natural world, as we invited scientists to consider what makes ecosystems on land and water resilient, and explored how organisms as diverse as tobacco plants and bacteria have evolved resilience strategies that offer lessons—or at least metaphors—we humans could embrace.

One psychologist coined the term ordinary magic to describe the mix of features that brews resilience. We learned that no clear-cut recipe exists. Investigating it carries pitfalls—in logistics, funding, and culture clashes. Despite the obstacles, scientists are parsing key ingredients of resilience. Some carry policy prescriptions; others require an uncomfortable rethinking of adaptations needed to survive. Perhaps most important, we found that this research cultivates something that's in short supply these days: It breeds hope.

AFTER THE DELUGE

Twelve years after Hurricane Katrina, social scientists seek lessons from its survivors

By **Kelly Servick**, *in New Orleans, Louisiana*

A muggy quiet has settled over New Orleans, Louisiana's Gentilly neighborhood as it soaks up a late-September rainstorm. Deep puddles hide dips in the street. And in a soggy patch of grass, a wooden kiosk tells a story of catastrophe.

"This place is a memorial to the trauma of the Flood," reads the text, written by a local nonprofit, Levees.org. Near here, a section of concrete levee gave way one August morning in 2005, sending the floodwaters of Hurricane Katrina crashing into the neighborhood. Yet the monument is not only a reminder of suffering, but also, the text insists, "a symbol of the residents' resilience and determination to return home."

Resilience and rebuilding—those two appealing themes bring hope after a natural disaster. The reality is more complicated. Many who fled Katrina's destruction never did return home. More than 12 years later, tidy brick houses in Gentilly are interspersed with empty lots while post-Katrina lives play out elsewhere.

Some of those survivors, wherever they ultimately ended up, are proving more resilient than others. "One household or family manages to recover," says David Abramson, a public health researcher who studies disasters at New York University in New York City. "The other remains dysfunctional."

Abramson has been surveying people affected by Katrina every few years since the storm. Poor, predominantly black families on cheaper property in lower-lying areas faced disproportionate damage from Katrina—and a harder road to recovery. But with the passage of years, the paths of survivors have diverged in complex, hard-to-predict ways. "Initially, I thought that those with the least would do the worst," Abramson says. "That wasn't always the case."

Abramson is one of three social scientists leading a project called Katrina@10. It's looking for long-term predictors of resilience—factors that cushion the shock of disaster and set the course for recovery. In their three long-running studies, the researchers have found a range of factors that seem to help, such as financial resources, social and cultural ties, and access to stable housing after the event, which all seem to help. Now, they're combining their cohorts to see whether those results will generalize. If the predictors they identify hold true across other natural disasters—and that remains to be seen—Katrina@10 could help policymakers and disaster recovery programs pick out especially vulnerable groups. It might even steer them toward interventions that do the most good.

Following survivors wherever they end up, year after year, is an unusual and costly proposition for a field in which di-

PHOTOS: (LEFT TO RIGHT) VINCENT LAFORET/POOL/AFP/GETTY IMAGES; MARIO TAMA/GETTY IMAGES

Workers clean a New Orleans, Louisiana, school (right) after Hurricane Katrina breached levees (left). Long after debris was cleared, families struggled to recover.

saster experts tend to lurch from one catastrophe to the next. Last year alone saw flooding across Houston, Texas; wildfires in California; and a crushing hurricane in Puerto Rico; to name a few. But studying survivors long after the floodwaters recede can pay off, the researchers say. "The 10- to 15-year time frame allows us to see what's real recovery," Abramson says, "and not just fleeting."

KATRINA SLAMMED into the Louisiana coast on 29 August 2005, and 80% of New Orleans was soon underwater. The city's Superdome, normally home to raucous football games, overflowed with refugees. Some families trudged out of the city on foot; others who couldn't escape waved for help from rooftops. It's estimated that more than 1800 people died and that the damage exceeded $100 billion. The country had never seen anything like it.

Katrina "is a flash point in people's minds about how bad it could really be," says Jeffrey Hebert, a city planning expert who from 2014 to 2017 served as the city's first "chief resilience officer."

Despite his catchy title, Hebert acknowledges that resilience has many meanings, some easier to measure than others. Engineers may gauge a city's physical resilience by the strain a levee can bear. Pinpointing what makes a person or community resilient is harder. But by a stroke of luck, two social scientists who later became leaders of Katrina@10 were uniquely poised to try. That's because both had been following New Orleanians before the storm for unrelated studies, and so were able to pivot and compare subjects' lives before with what came after.

One was Mark VanLandingham, a sociologist at Tulane University here. In 2002, he launched a project in the quiet area of eastern New Orleans comparing the lives of Vietnamese immigrants who had settled here after evacuating from Saigon in 1975 with those of families who stayed behind in Vietnam. In the summer of 2005, his team was wrapping up a survey on the health and well-being of people in 125 Vietnamese households.

Meanwhile, another sociologist, Mary Waters from Harvard University, was part of a nationwide study examining how higher education affects the health of single parents. The team had reached about 500 first-generation college students in the New Orleans area for a phone survey when Katrina sent them fleeing for dry ground.

Waters, safe and dry in Cambridge, Massachusetts, and VanLandingham, who es-

caped to Galveston, Texas, before his own house took on a meter of water, didn't know each other. They didn't know much about disaster research. But both immediately recognized that their questionnaires documenting the health, social networks, and personality traits of Vietnamese immigrants and mostly poor, black, single mothers before the hurricane had taken on outsize significance.

In the months after Katrina, Waters and VanLandingham, along with their colleagues, began tracking down their displaced participants to see how they were faring. The researchers tried calling the phone numbers on file and sent teams to search New Orleans neighborhoods for participants or friends who might know where to find them.

Meanwhile, Katrina's devastation also drew Abramson in. He had been exploring the impact of HIV/AIDS in New York City, but the storm inspired him to lead a caravan of about 30 researchers, graduate students, and health workers to visit temporary housing sponsored by the Federal Emergency Management Agency (FEMA) in Mississippi and Louisiana. Their goal was to monitor those families over the coming years as they sought permanent housing back in their original neighborhoods or elsewhere, and to track how disaster and displacement affected health.

In a first round of surveys, Abramson's Gulf Coast Child and Family Health Study interviewed people from 1079 displaced households between 6 and 12 months after the storm. As the team's 12-passenger vans rolled through FEMA housing sites, they found families of six crammed into trailers, uncertain whether they'd be forced to move out on a few days' notice. Some feared for their safety and kept their children inside. "It made for a very claustrophobic and depressing situation," Abramson says.

Abramson would track those families over time and watch their paths diverge. But in another population, a future colleague of VanLandingham's saw a different trajectory from the start. Cam Tran had immigrated from Vietnam as a child, and after Katrina she traveled from her home in Texas to New Orleans to help her in-laws recover. Tran remembers the day she drove into their neighborhood, about a month after the storm.

"It was completely gloomy and dark," she says. "No sound." But as Tran approached Mary Queen of Vietnam Church, she heard music from a car radio and saw neighbors rebuilding the church roof. "We asked them, 'Is it safe for people to come back?' and they said, 'Well, you know, there's no electricity or water or anything like that. But yes, please do come back!'"

Tran took their advice. She moved here and helped set up a charter school. And she later became a coordinator for VanLandingham's study, Katrina Impacts on Vietnamese Americans in New Orleans, which showed that the optimistic welcome she received from the rebuilders presaged an entire community's long-term recovery. In the coming months, VanLandingham watched members of the community wake at daybreak, drive back to their neighborhood, and rebuild—one house at a time. They seemed to embody resilience.

Two years later, when VanLandingham and Abramson met for the first time at a conference here, they discovered that some of their participants hailed from adjacent neighborhoods. Together, as the pair drove those streets in VanLandingham's Subaru Outback, something started to click: Families in the two studies had similar economic means and their homes had sustained similar levels of damage. Conventional wisdom might have predicted similar recoveries. But it was "as if they had almost suffered two entirely different events," Abramson says.

The neighborhood of Abramson's mostly black participants, the ones who'd wound up in FEMA housing and whom Abramson was now carefully tracking, was still strewn with debris and abandoned belongings. In a preliminary analysis, that group was scoring well below VanLandingham's Vietnamese families in mental health surveys. Why did such gaps exist between those communities when it came to resilience, the researchers wondered, and could anything be done to narrow them?

Sociologist Mark VanLandingham visits Mary Queen of Vietnam Church in the New Orleans, Louisiana, community he's still studying post-Katrina.

PHOTO: © WILLIAM WIDMER 2018

YEARS PASSED, but the sociologists didn't leave. For Waters, there never seemed to be a good time to stop. "We didn't set it up to be a study that was going to last 10 or 15 years," she says. But in each round of interviews, "it was so clear that we were in the middle of the story."

By 2009, the women in Waters's Resilience in Survivors of Katrina (RISK) Project were scattered across 23 states, and just 16% had returned to their prehurricane homes. The RISK researchers examined mental health trajectories, in particular whether those women had returned to their level of psychological functioning from before the storm. Some had, among them "Keanna," who built a new life in Houston with her husband and five children. She re-enrolled in school and started her own business; she said she had developed a deeper relationship with God. On the other end of the spectrum sat "Belinda," also a mother of five, who spent nearly a year at a friend's house in Arkansas before returning to New Orleans. She became estranged from her partner, struggled to support two unemployed sisters, and faced depression and weight gain.

Some of the factors widening that divide were predictable. In the RISK Project, researchers found that stressors such as going without food or water after the storm or, worse, losing a loved one predicted longer-term mental health struggles, as did reporting a weak social support network before Katrina. But other findings took Waters by surprise—such as the fact that, controlling

for all other factors, the loss of a pet because of the storm had lasting negative effects.

Abramson, meanwhile, developed an analytical tool to gauge recovery on the basis of measurements in five areas: physical and mental health, economic stability, stable housing, and "social role adaptation," or how people feel that they fit into their community. That framework allowed him to identify prestorm factors that most contributed to long-term recovery. For example, measures of "psychological strength"—which included religiosity and the perceived ability to adapt to stressors—were most predictive of a strong recovery. Having a household income of at least $20,000 was close behind. Being older than 50 or disabled had strongly negative effects on recovery, as did spending a prolonged period displaced from one's home. How returning home versus resettling elsewhere influenced recovery remains an open question.

VanLandingham's study took yet another tack: It became a deep dive into the role of culture and history in resilience. Interviews with some of his original study participants and community leaders suggested that the shared experience of the Vietnam War and immigration had united neighbors, motivating them to rebuild. In a book published last year, *Weathering Katrina: Culture and Recovery among Vietnamese-Americans*, VanLandingham also suggested that this community's members bounced back faster than many black residents of similar means because they faced less discrimination.

LONG-TERM RESILIENCE studies such as these are unusual, in part because funding for them is hard to sustain. And in 2012, VanLandingham's prospects for continuing his project looked bleak. His application for new funding from the U.S. National Institutes of Health (NIH) was rejected. Reviewers were mostly positive, but they complained that he had no comparison group, no way to put his findings into context. Then, an NIH program officer told him that he wasn't alone.

"She said, 'There's this woman at Harvard having the same problem,'" VanLandingham recalls. He contacted Waters, and they recruited Abramson. In 2015, the trio won about $6 million in NIH funding over 5 years for what was finally, a decade after the storm, a unified effort: Katrina@10.

The study has an ambitious goal: to build a crystal ball that uses a few characteristics to predict disaster recovery in the long term. The effort includes a new round of standardized surveys in the three original cohorts, plus two other data sets to put them in a broader context. One data set is from the U.S. Census Bureau and covers New Orleans's changing demographics. The other is drawn from a random sampling of people who had lived there before Katrina and includes information on health and well-being. The study's findings could help other communities traumatized by fires, floods, and earthquakes, by identifying people at highest risk and how best to help them.

Abramson already has a hunch about one factor that will rise to the top, on the basis of unpublished data from his cohort, which started off in those FEMA trailers. "The faster you move somebody into stable housing, the faster, more accelerated, and more durable their recovery will be," he predicts. If he can confirm that suspicion in the larger Katrina@10 cohort, it could help improve how emergency response agencies operate. For example, recovery programs could invest in more durable housing for evacuees rather than provisional camps, he says.

But the researchers also come back to what they've seen firsthand: Different communities have different needs, and different strengths and weaknesses. Abramson envisions a future in which organizations that step in to help after a disaster can gauge how resilient the person sitting in front of them is likely to be.

For now, Katrina@10 has a more prosaic task at hand: rounding up its megacohort of the roughly 2200 participants from the three original studies for one last interview. A team of graduate students has helped track participants online when the numbers and contacts on file led nowhere. One student found a participant by tracing the auto body shop uniform he was wearing in a Facebook photo.

In the most recent round of interviews, some participants seemed befuddled that the researchers were still at it. But Tran noticed a change in their attitudes after Hurricane Harvey struck Houston, a city that welcomed many New Orleans refugees in 2005. Harvey's landfall last summer, almost exactly 12 years after Katrina's, brought back memories—and nurtured a grim camaraderie. "It was like, 'Oh my gosh, now we've got to find some way to help the Houston community because of what they did for us,'" Tran says.

Abramson is plotting studies of resilience in Hurricane Harvey survivors—along with people coping with the aftermath of Hurricane Maria, which hit Puerto Rico weeks later—to compare their trajectories to what he's seen in Katrina survivors. If common drivers of resilience emerge across varied disasters, Katrina@10 participants may end up helping fellow survivors in more ways than they ever imagined. ■

PHOTO: D. P. WILSON/FLPA/SCIENCE SOURCE

NATURE'S STRATEGIES

Fish that switch sex to thrive

Fish are masters of reproductive resilience. About 450 species switch sexes over their lifetimes to maximize their number of offspring. The fish do so by undergoing hormonal changes that transform their organs from those of one sex to the other. Patterns of sex switching vary by species. Big females produce more eggs than little females, so for some species, such as clownfish, it's best to be a male early in life when more runty and then switch to a female later on. But among males that fight each other for females or territories—such as groupers, sea breams (pictured above), and porgies—being a too-small male can mean no offspring at all. In that case, it's better to stay a small female instead.

Now, this age-old strategy is allowing fish like the sea bream to adapt to a modern challenge that also disrupts the sex balance: overfishing. Fishers favor the biggest catch. Because one sex is usually bigger than the other, the bigger sex risks being fished out. But researchers have found that sea breams—flavorful, reddish fish common in warmer Atlantic Ocean coastal waters—are ready. Removing big males prompts earlier-than-usual sex changes in some females, so the sex balance is preserved. Still, it's more a short-term strategy than a long-term solution, researchers say. The fish are switching sex at younger ages, so females don't have a chance to grow big. That trend translates into fewer offspring and a shrinking population. That resilience strategy keeps them reproducing for now—but the fish can't save themselves all on their own. —*Elizabeth Pennisi*

LESSONS IN RESILIENCE

In war zones and refugee camps, researchers are putting resilience interventions to the test

By **Emily Underwood**

In 2015, in the name of science, more than 800 teenage boys and girls in northern Jordan each allowed 100 strands of hair to be snipped from the crowns of their heads. Roughly half the teens were Syrian refugees, the other half Jordanians living in the area. The hair, molecular biologist Rana Dajani explained to the youngsters, would act as a biological diary. Chemicals embedded inside would document the teens' stress levels before and after a program designed to increase psychological resilience.

It was a unique experiment. And it was one that suited Dajani, who's based at The Hashemite University in Az-Zarqa, Jordan. Dajani looks askance at many humanitarian interventions imported from elsewhere. "I'm always skeptical of any program coming in from the outside, which says they can heal or help," she says. Half-Syrian herself—Dajani's mother is from Aleppo, her father from Palestine—she was also eager to study the physiological effects of conflict. So when medical anthropologist Catherine Panter-Brick, whom Dajani had met at Yale University in 2012, approached her about putting the resilience-boosting program to the test, she seized the opportunity.

Run by the nongovernmental organization (NGO) Mercy Corps, headquartered in Portland, Oregon, and Edinburgh, the Youth Take Initiative—or, in Arabic, Nubader program—would teach stress management and relationship skills to at-risk 11- to 18-year-olds. Nubader falls into a booming category called psychosocial support; the interventions are as diverse as play therapy, parenting courses, and mindfulness training, and they've flourished across more than a dozen countries. Many aim to enhance the resilience of children affected by war and other disasters.

Smoke rises from a November 2017 airstrike in Damascus carried out by the Syrian government. Since the conflict began, millions have fled the country.

Finding ways to support these children has never been more urgent. Hundreds of millions of young people live in countries riven by armed conflict. Roughly 15% to 20% may develop posttraumatic stress disorder (PTSD) and other mental illnesses. Psychosocial programs, usually staffed by laypeople with various levels of training, are feasible in war zones and refugee camps in a way that specialized psychological care often is not. The question is: Do they work?

That's where the hair collection came in. Panter-Brick and Dajani hired professional hairdressers, who collected the strands while offering the teens stylish hairdos. The samples were then shipped to a lab at the University of Western Ontario in London, Canada. While the Canadian scientists ground up the strands and measured levels of the stress hormone cortisol, research assistants interviewed the teens about past traumas and current stress.

On average, the Syrian cohort reported six traumatic experiences, most commonly witnessing bombardments and having their homes forcibly searched or demolished. As Dajani listened to their harrowing stories, she wondered whether Nubader's setup, just 16 sessions of psychological coaching, had the power to deliver on the nonprofit's ambitious goal: boosting resilience by alleviating stress, strengthening relationships, and "healing the scars of conflict."

THE STUDY OF PSYCHOLOGICAL RESILIENCE has its roots in the 1970s. That's when Norman Garmezy, a developmental psychologist at the University of Minnesota in Minneapolis, began studying schoolchildren who thrived despite severe hardship, such as neighborhood violence or parents with mental illness. After Garmezy retired, his students picked up where he left off, pinpointing factors that helped these children cope. Some were environmental, such as a

PHOTO: DIAA AL DIN/ANADOLU AGENCY/GETTY IMAGES

A young Syrian refugee (right) who fled to Jordan listens to a teacher (left) as part of a Mercy Corps youth program.

PHOTO: SEAN SHERIDAN/MERCY CORPS

strong bond with a parent. Others bloomed from within, such as a sense of agency or control over one's fate. One of Garmezy's students, developmental psychologist Ann Masten, coined a term for the constellation of variables that together help a child transcend bad circumstances: ordinary magic.

What began with Garmezy and the resilient children in urban Minneapolis raised an obvious question: Can resilience be taught to others who might not come by it as easily? Or, put differently, can ordinary magic be brewed for just about anyone?

Before answering that question, social scientists and psychologists had to consider what, exactly, resilience is. They have yet to agree. Some believe resilience means restoring mental health after a traumatic event. Others consider it a conscious determination to persevere under difficult circumstances. Still others describe it as a child's ability to benefit from external resources, such as a caring adult. To complicate matters, humanitarian groups use the term resilience to describe any or all of these positive outcomes.

"It's quite a squishy concept," says Jon Kurtz, Mercy Corps's director for research and learning in Washington, D.C. By collaborating with Dajani and Panter-Brick, Mercy Corps hoped to get a firmer grasp on how to support and measure resilience in the Syrian and Jordanian teenagers, he says.

Despite the cacophony of definitions, most studies of resilience interventions in children ask one of two questions: Does a program promote existing mental health by helping children cope with war and displacement? Or does it prevent mental health complications for which children are now at higher risk?

Outcomes are mixed for the few resilience programs that scientists have evaluated. In 2016, an article in *Current Psychiatry Reports* reviewed data on 24 mental health and psychosocial programs conducted in nine countries, including Bosnia, Uganda, and Nepal. The researchers found that although all interventions had some positive impact on mental health, less than half met their goals. Nearly a quarter had a negative impact on an endpoint the program aimed to improve, such as symptoms of depression or PTSD. Some programs worked in one country but failed in another: Teaching emotional regulation to former child soldiers in Sierra Leone improved their social relationships, for example, whereas a similar effort for Palestinian children increased symptoms of PTSD. In a study in Nepal, being affiliated with a political movement appeared to protect mental health among former child soldiers. Yet the opposite was true among children in Bosnia, says Wietse Tol, one author of the 2016 review and a mental health researcher at Johns Hopkins University in Baltimore, Maryland.

What accounts for these inconsistent outcomes? The factors that support mental health and resilience in one situation may be useless or even harmful in another, Tol says.

To pinpoint the ideal interventions for a community, researchers need to spend time there, suggests Michael Pluess, a psychologist at Queen Mary University of London. In recent focus groups with Syrian refugee mothers in Lebanon, for example, Pluess and his colleagues found that a popular ingredient in many psychosocial programs—a concept called "internal locus of control"—was problematic among people anchored by religion. An internal locus of control is the conviction that success comes thanks to one's own efforts, such as hard work, rather than external factors. Although often seen as supporting mental health, the concept didn't resonate with religious parents who believe that life unfolds according to God's will, Pluess says.

Despite the mixed results of resilience programs, Tol is heartened by the learning curve he sees. "I think the research is showing that it is possible to teach resilience" to conflict-affected children, he says.

A MEDICAL ANTHROPOLOGIST at Yale, Panter-Brick has navigated that learning curve herself, traveling extensively to study resilience. She has visited Nepal and interviewed homeless children. In Afghanistan, she probed the mental health of young people in the wake of war. Panter-Brick argues that for children, resilience has three dimensions: individual strengths, relationships with family and peers, and community support.

Nubader mainly targets the first. The program nurtures an adolescent's resources and skills, although it also aims to build a support network for teens by training mentors and creating community councils to consider children's ongoing needs. Between 2014 and 2016, more than 4000 young people with mental health struggles and poor access to social services participated in similar Mercy Corps–run programs across the Middle East.

The intensive program Panter-Brick and Dajani evaluated in Jordan lasted 2 months.

> ***"It takes guts to let someone in to evaluate your program as thoroughly as we did."***
>
> **Catherine Panter-Brick**, Yale University

In it, teenagers gathered at a youth center twice a week to participate in group activities of their choosing, including soccer, sewing, and computer repair. Those activities were meant to foster social bonds and build confidence and competence. Participants also learned how chronic stress can affect the brain—for example, by impairing impulse control. Coaches practiced relationship-building skills with the teenagers, such as expressing affection and empathy.

This neuroscience-based instruction, called the Profound Stress and Attunement model, was developed by former Mercy Corps youth program director Jane MacPhail. It's loosely based on emerging neurobiology research suggesting that social relationships can buffer the negative effects of chronic stress and trauma.

From its own before-and-after program evaluations, Mercy Corps believed that MacPhail's program worked. But those evaluations lacked the scientific rigor of an independently run randomized trial, which would compare the intervention with another activity or with no intervention at all. So the NGO approached Panter-Brick and Dajani for an outside assessment. "It takes guts to let someone in to evaluate your program as thoroughly as we did," Panter-Brick says.

Accomplishing that goal meant running Nubader and testing it simultaneously. Panter-Brick and Dajani invited 817 young people living in Jordan, who had already signed up for Nubader, to participate. They included both Syrian refugees and at-risk Jordanian teens. The youths were randomly assigned to the program or to a 2-month waitlist, which served as a control group.

Dajani and Panter-Brick quickly found that their desire for rigor and clear outcomes ran up against teenage inhibitions and logistical snags. Plans for an expansive collection of biological samples, including cheek swabs for DNA, dried blood spots to test immune function, and saliva for additional cortisol levels, had to be pared back. Many teens were too embarrassed to spit into the sample vials. Saliva was also hard to freeze and transport in Jordan's summer heat because electricity was sporadic, Panter-Brick says.

Hair could be mailed cheaply in an envelope. Still, even gathering those samples "was very tough" at first because of the tense relationship between Jordanians and the Syrian refugees they were hosting, says Natasha Shawarib, the project manager for Mercy Corps in Amman. Some refugee families feared their child's data would be handed off to the Jordanian government, and they

NATURE'S STRATEGIES

Resilience by regeneration

Humans should envy the axolotl (pictured, right). Our powers of regeneration are limited: Broken bones knit, wounds heal, and large parts of the liver can regenerate, but that's about it. But the axolotl—a large salamander also called the Mexican walking fish because it looks like a 20-centimeter eel with stumpy legs—can replace an entire missing limb or even its tail, which means regrowing the spinal cord, backbone, and muscles. About 30 research teams are probing how these salamanders do it. In the axolotl, they've found, various tissues work together to detect limb loss and coordinate regrowth. In the process, the animals reactivate the same genetic circuits that guided the formation of those structures during embryonic development, causing generalist stem cells to specialize.

Axolotls are only one of several regenerators in the animal kingdom. Flatworms called planarians are even more resilient—able to surge back after losing 90% of their bodies. One small fragment of those 2-centimeter-long aquatic worms can rejuvenate the brain, skin, gut, and all the other functional organs. Again, stem

cells are key, and a special set of genes active in muscles tells those stem cells what to do, activating growth and specialization genes in the right cells at the right time. So the planarian can rebuild itself almost from scratch, whereas the axolotl can rebuild only if the main body axis is intact. This year, researchers took another step toward detailing the molecules underlying regeneration by sequencing the genomes of those two species. The ultimate hope: One day, we'll be able to coax injured humans to execute the same repairs. —*Elizabeth Pennisi*

PHOTO: DANTÉ FENOLIO/SCIENCE SOURCE

needed reassurance that the information was purely for research, she says.

Before and immediately after the intervention, and again 11 months later, the teenagers also answered surveys about their mental health and sense of security. "To what extent do you fear for your family in your daily life?" one question asked. "To what extent do you fear or worry about losing your family's source of income?" read another.

To lighten the mood, one of Dajani's research assistants—a Syrian refugee herself—volunteered to paint the girls' nails before interviews. But those nail painting sessions often ended in tears. Both the fieldworkers and the teenagers came to dread the questionnaires because talking about past traumas was so upsetting, Dajani says.

In response to the teens' and fieldworkers' requests, "we decided to ask the teenagers how they deal with negativity, not just remind them of it," Dajani says. The team crafted an Arabic translation of a survey called the Child and Youth Resilience Measure, originally developed by Canadian psychologist Michael Ungar. It queried the teens about sources of resilience in their own lives by asking them to rate 12 statements such as "I am aware of my own strengths," and "My family stands by me in difficult times," gauging their feelings of belonging and optimism.

In the end, the scientists determined that Nubader had a positive impact—but whether it nurtured resilience depends on whom you ask. The teenagers enrolled in Nubader felt moderately safer and more secure than members of the waitlisted control group—a benefit sustained 11 months later, the team reported in *The Journal of Child Psychology and Psychiatry* in October 2017. Findings from the hair strands, too, suggested a benefit: Average cortisol levels in the intervention group dropped by a third, the researchers reported in January in *Psychoneuroendocrinology*. In a subgroup with statistically low levels of cortisol—a phenomenon linked to higher risk of PTSD—cortisol production increased by nearly 60%, a healthy sign.

Those changes aren't dramatic. But participating in the program was clearly better than nothing, says Danny Pine, a child psychiatrist at the U.S. National Institutes of Health in Bethesda, Maryland.

Mercy Corps interpreted the findings as a win. "Now we can confidently say that our work does make a difference," says Noura Shahed, a project coordinator at Mercy Corps in Amman. But Dajani and Panter-Brick say the reality is more nuanced: Although teens had less fear and stress, the study did not meet the scientists' strict definition of resilience, and the program did not appear to strengthen teens' social support, even though Mercy Corps's internal evaluations suggested that it did.

Dajani and Panter-Brick suspect that's because the intervention lasted just 8 weeks and largely targeted one source of resilience—individual strengths. "You could go to a great program every day, but if you go home and your family life is terrible, you're not going to build resilience," Dajani says. The surveys support that conclusion: Teens who scored high on the resilience measure from the start described close family ties and supportive communities, Panter-Brick says.

Resilience "isn't simply in the child, but embedded in their family, caregivers, and community," agrees Masten, who has noticed that same trend among children in Minnesota and elsewhere. That doesn't mean Nubader didn't benefit the teenagers in Jordan, Panter-Brick says. But ideally, she says, interventions should be more sweeping, reaching parents and communities, too. Mercy Corps is now doing just that through a support program for parents and caregivers that teaches them about the impacts of long-term stress on the brain, Shahed says.

Dajani and Panter-Brick's experiment was "important, even groundbreaking," Pine says. The experiment wasn't perfect, in part because the control group was just a waitlist. Comparing it with a slightly different program—for example, recreation with no educational content—would have helped the researchers identify the active ingredient of Nubader's success. Still, Pine says, the team showed that rigorously testing humanitarian programs under trying circumstances is possible.

Other NGOs are now applying science to their resilience interventions. In Lebanon, War Child Holland, a branch of the global NGO that assists children in conflict zones, is evaluating three efforts: a life skills program, a program to reduce parents' stress, and a World Health Organization–designed mental health intervention for Syrian refugees. War Child Holland's ultimate goal is to find the best way to support resilience at the individual, family, and community levels all at once, says psychologist Mark Jordans, War Child Holland's director of research and development in Amsterdam.

For Panter-Brick, one of the most valuable lessons out of Jordan came from the young people themselves. They reminded her that resilience research "is not about rescuing victims of chaos," she says. Rather, it calls for identifying potential sources of strength that young people can draw on to survive, even thrive. "It's about reshaping your lens on the world," she says, "to what people feel respects their dignity." ■

Reporting for this project was supported by a Rosalynn Carter Fellowship for Mental Health Journalism.

NATURE'S STRATEGIES

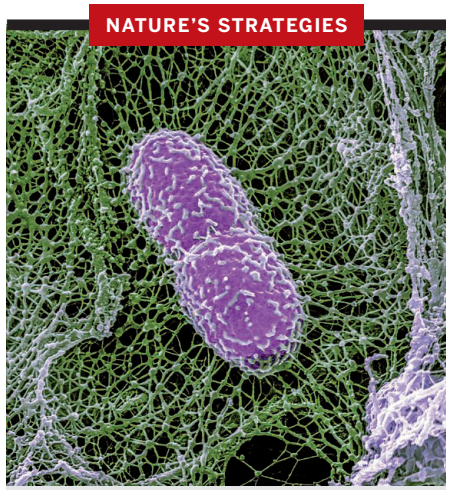


PHOTO: SPL/SCIENCE SOURCE

Stealing genes to survive

Imagine a raging infection in the lungs of a hospitalized cancer patient. When a powerful antibiotic floods the patient's system, the bacterium responsible, *Klebsiella pneumoniae* (pictured above), seems to be doomed. But it can deploy a resilience strategy honed over billions of years: borrowing a gene from another cell that enables the pathogen to survive.

When environments change, organisms adapt or die. *K. pneumoniae* and other bacteria have turbocharged the process of adaptation by snagging genes from elsewhere, including various bacteria and DNA molecules loose in the environment. Such horizontal gene transfers allow the bugs to gain valuable new traits, everything from the ability to thrive on cheese rinds to antibiotic resistance.

Researchers think that *K. pneumoniae* acquired its antibiotic disrupter gene, *blaKPC*, from another, still-unidentified bacterium. Bacteria outfitted with the gene churn out an enzyme that breaks down several antibiotics.

As with many resilience strategies in nature, stealing genes has its costs. Sometimes microbes incorporate harmful genes instead of helpful ones. And much like a new player on a basketball team, the protein produced from an acquired gene may not mesh with the cell's other proteins. But unfortunately for patients, *K. pneumoniae*'s strategy works all too well: Those bugs kill between 40% and 70% of the people they infect. —*Mitch Leslie*

BENDING TO THE WATER'S WILL

In flood-prone Bangladesh, resilience can mean letting water have its way

By **Warren Cornwall**, *on Polder 32 in Bangladesh; Photography by* **Tanmoy Bhaduri**

For Jaharul Sardar, a rice farmer in rural Bangladesh, the perils of living behind a wall hit home one cloudy May afternoon in 2009. Sardar was standing beside his fields when he heard neighbors cry out in alarm. A black hill of seawater was sweeping toward him. Sardar's wife and 5-year-old son clambered atop the embankment that stands guard over his mud-floored house. Instead of joining them, he dashed inside to rescue a suitcase stuffed with cash and property records.

Within seconds, waves slammed into the house. Sardar was trapped. Water pushed the structure, with him inside it, into an adjacent pond. His wife pulled him to safety, soaked but alive.

For Sardar, relief was short-lived. That day marked the beginning of fierce floods. Each day the tide poured in through the breached wall, drowning fields and homes in saltwater, and then withdrew and left a blanket of mud. "During high tide it was all underwater. And during low tide it was like [a] desert," Sardar says.

The 4-meter-high earthen wall that failed dated from the 1960s. It encircles 80 square kilometers of land to create a massive "polder"—an artificial island surrounded by the vast tidal rivers that extend like thick ten-

Villagers cluster on Polder 32, an artificial island in southwest Bangladesh with an uncertain future.

drils from the nearby Bay of Bengal. At 66, Sardar remembers conditions before the wall, when flooding during very high tides was the norm here. Until that fateful day in 2009, the flooding appeared to be banished.

Bangladesh, a vast river delta that barely rises above the sea at the best of times, is buffeted by natural forces including flooding rivers and cyclones blowing in from the bay. Over decades, the country has developed defenses: warning systems, storm shelters, salt-resistant crops, and 139 polders near the coast—a 5700-kilometer network of walls to protect farmland from inundation. But humanmade infrastructure is not infallible and can cause problems of its own. That's starkly apparent across the country's polders, which have disrupted a fragile standoff between water and land and are now straining to hold back the water. As climate change compounds that threat with rising seas and stronger storms, Bangladeshis who have spent years building barricades are considering what was once unthinkable: letting the water in. It's resilience by bending, not resisting. And it's tougher to do than it sounds.

Other countries are trying similar approaches. Vietnam recently adopted plans to allow more flooding in the upper reaches of the Mekong delta. The Netherlands, renowned for building some of the world's most sophisticated sea walls, is adapting suburbs for controlled river flooding. On the other end of the spectrum sits Indonesia, which is planning a $40 billion, 40-kilometer-long sea wall to shield its capital city, Jakarta, from the Java Sea.

Here in a country the size of Iowa that's home to 165 million people, experiments in resilience underscore the challenges of rearranging a crowded landscape. "You can't remove the polders now," says Anisul Haque, a hydraulic modeler at the Bangladesh University of Engineering and Technology (BUET) in Dhaka who studies coastal flooding. "So what can you do?"

Polder 32 resident Jaharul Sardar, who narrowly escaped a 2009 flood, remembers when the walls ringing his home kept him safe.

RIVERS ARE THE MIDWIVES of Bangladesh. The Ganges and Brahmaputra pour from the Himalayas and converge with the Meghna River to form the world's fourth largest drainage, which flows into the Bay of Bengal. Monsoon rains routinely put a quarter of the country underwater. The flooding brings hardship, but it also nurtures the rice that feeds one of the most densely populated nations on Earth.

The country itself is born from those rivers. An estimated 1 billion tons of sand and silt flow downstream every year and settles in the delta, counteracting relentless erosion. Geologically, Bangladesh is a giant sandbox, 90 meters deep in places.

Ainun Nishat knows those rivers with an intimacy earned by spending 4 decades studying them. Sitting in his third-floor office at BRAC University's Centre for Climate Change and Environmental Research in Dhaka, the engineer describes how good intentions have brought unexpected consequences.

Polders—the name is borrowed from the Dutch, who used a similar strategy to carve farmland from marshes—were first built in Bangladesh in the 1960s. But although polders allow more intensive farming, Nishat says, "They are also a problem." The walls impede the natural movement of water and sediment. Rivers now funneled between artificial embankments are filling with silt. Land inside the polders, starved of new soil that would otherwise flow in, is sinking. Polders are turning into bathtubs that, if something goes wrong, can fill with water.

Meanwhile, sea level is projected to rise 0.4 to 1.5 meters on the Bangladesh coast by 2100. Episodes of extremely high water driven by storms and tides, which today occur once a decade, will probably happen

NATURE'S STRATEGIES

Squirrels with a rainy day fund

Scurrying around the South Dakota prairie, 13-lined ground squirrels (pictured, right) mark the approach of winter by bingeing. By the time a squirrel holes up to hibernate, its weight will have soared by about 40%, thanks to extra fat that will tide the creature over until spring.

During droughts, migrations, bleak winters, and other challenges, organisms often face times when resources are scarce. To get by, the ground squirrel, like many other creatures, stockpiles resources to use later. It can gain more than 2% of its body weight in a single day as it gorges on seeds, grasshoppers, and other delicacies.

But the tactic has downsides. A roly-poly rodent is easier prey for a hawk or coyote. The rainy day fund also can run out prematurely. So once a squirrel is nice and tubby, it enters hibernation, slashing its energy expenditure by 90%. Its body temperature drops to just above freezing and its heart rate falls to as low as 5 beats per minute, down from the usual 350 to 400.

Packing on the fat requires metabolic and behavioral adjustments. But somehow, the squirrel dodges the health problems that plague obese people. Although it develops some of the metabolic defects of type 2 diabetes, the animal isn't sick. And by spring, it is lean and spry and ready to begin the cycle again. —*Mitch Leslie*

PHOTOS: (TOP TO BOTTOM) TANMOY BHADURI; GRACHEVA LAB

three to 15 times every year at the end of the century, according to a 2015 study by U.K. and Bangladeshi researchers. That trend will put the polders and their inhabitants at even greater risk.

The devastation of Sardar's polder, Polder 32, starkly illustrates the dangers posed by that confluence of climate change and decades of hydraulic tinkering. That day in 2009, a wave of water originating from nearby Cyclone Aila combined with strong currents to burst through embankments on several polders. The disaster left more than 150 dead and $270 million in damages in Bangladesh.

Many believed the answer was to strengthen the polders. In 2013, the World Bank committed $400 million to raise embankments on 17 polders that are home to nearly 800,000 people. It's a first step in a push by the Bangladeshi government to build up the entire polder system with an eye toward rising seas.

What Steve Goodbred, a coastal geologist at Vanderbilt University in Nashville, saw when he visited Polder 32 after the storm pointed to a different approach. Goodbred and a research team found that land inside the polder was more than a meter below the average high tide in the area. As long as the walls held up, the sinking went largely unremarked upon. But the cyclone laid bare the risks. "It was, 'Wow, no wonder the flooding was so bad,'" says Goodbred, who has spent the past 2 decades studying the Bangladesh delta.

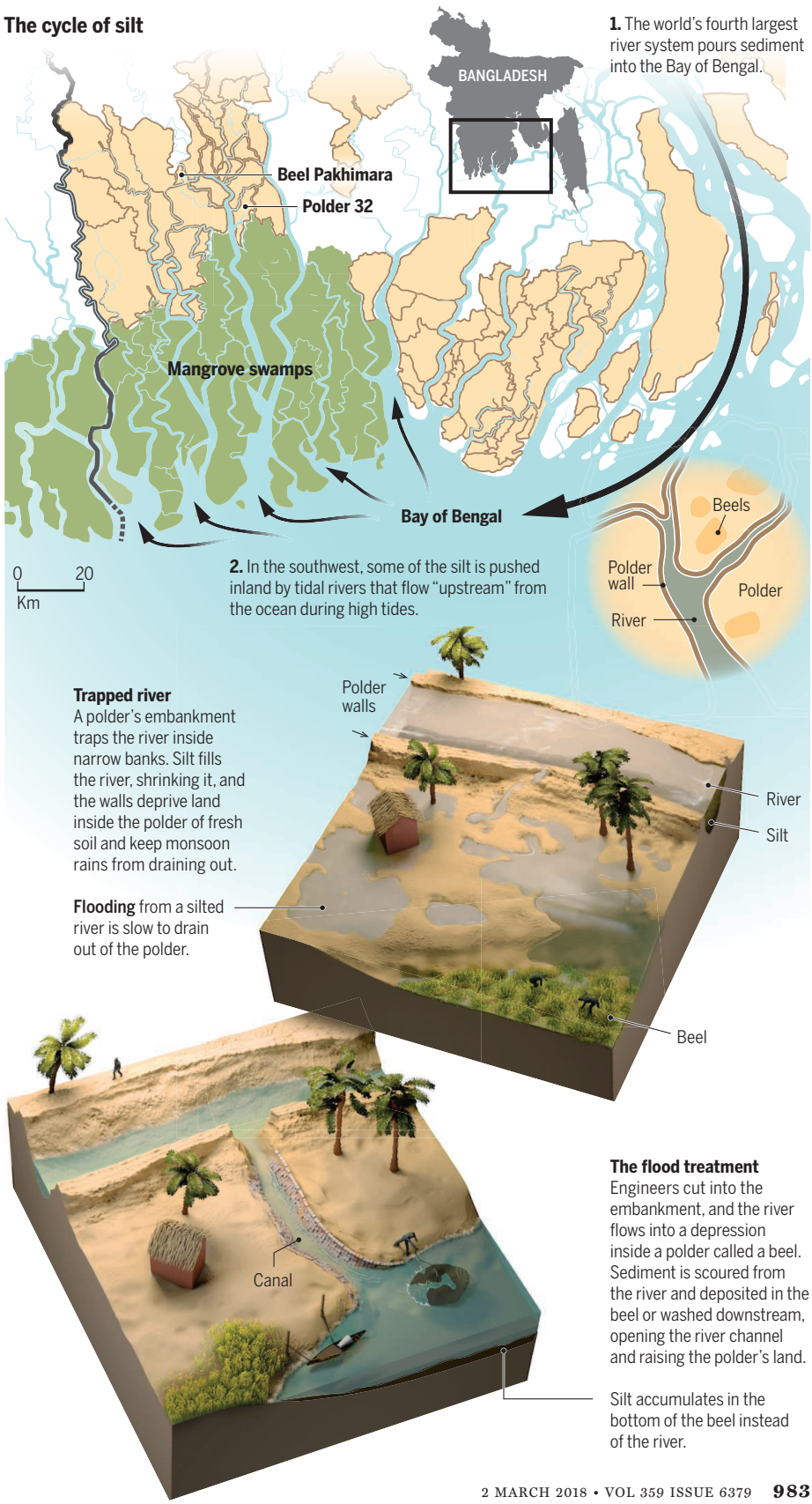
However, Goodbred also found cause for optimism. Land just outside the polder was 10 centimeters above the typical high tide, suggesting that, without human intervention, natural sediment deposition would keep the land above all but the highest tides. Furthermore, the disastrous flooding here had an upside, delivering enough new silt to raise land inside the walls more than a third of a meter on average.

That outcome led Goodbred to suggest letting some water in. In the short term, controlled flooding into the polders might be painful. In the long run, it could elevate land and minimize damage from future breaches. In a world of rising sea levels and bigger storms, Goodbred says, "I think that has to be a part of any long-term solution."

IT'S A SEDUCTIVE IDEA. It's also not a popular one for many who would have to live with the consequences, including government officials. "We don't need to raise the land. Farmers are not demanding it," says Abdul Hannan, the top engineer in this region for the Bangladesh Water Development Board, the powerful government agency that manages the country's water.

Water meets land in Bangladesh

Many Bangladeshis live on artificial islands called polders, behind walls built to protect the low-lying islands from floods. But the polders disrupt sediment deposits and face climate change pressures, and it's harder than ever to keep water out. Residents are considering new ways to protect the polders and their land.



GRAPHIC: J. YOU AND V. ALTOUNIAN/*SCIENCE*

A woman walks by a canal linking the Kobadak River to a polder. Connecting the canal and the river causes flooding in one place that might protect land elsewhere.

PHOTO: TANMOY BHADURI

Yet about 30 kilometers inland, farmers spurred an experiment to do just that. It's much like the prescription Goodbred offered—use controlled flooding to raise sunken land and drain polders—but this experiment began almost by accident in the early 1990s.

These inland polders are designed to protect against flooding by seawater driven far up river channels during high tides. After the embankments were built, those rivers, with smaller, slower currents than rivers closer to the coast, started clogging with silt that would otherwise have settled on land. Meanwhile, the polders began to sink. After monsoon rains, water in the polders failed to drain through small canals into surrounding, now sluggish, rivers. Farms and towns inside stewed in stagnant water for months.

Local ingenuity and desperation prompted farmers on two river systems in southwest Bangladesh to cut gaps in embankments to open the polders to the river. The government cracked down with arrests. Then, something unexpected happened—exactly the outcome Goodbred later saw at Polder 32. Within a few years, the land inside the inland polders had risen a meter or more. The rivers grew deeper. Waterlogging eased.

Since then, government officials have tried to replicate the success of those first improvised projects. The latest test for what's known as "tidal river management," or TRM, began in 2015. A construction crew cut through the embankment of a polder lining the Kobadak River. Workers then dug a canal joining the river outside the polder to a 660-hectare depression inside it—a wetland known as a beel.

Visit today and the experiment's effects are evident. Parts of the beel, called Beel Pakhimara, have gained half a meter of fresh land. The river runs faster, as sediment is deposited in the beel instead of the riverbed. As the main river channel grows deeper, water is once again draining from nearby polders.

The endangered Ganges River dolphin has put in appearances since the flow returned, says Jahin Shams Sakkhar, an official with Uttaran, a local organization that has campaigned for projects like this one. "The more open it is, the more good it is for nature and for people," he says.

Yet the filling of the beel hasn't gone entirely as planned, partly because "so far, the whole TRM process has been based on assumptions," says Shah Alam Khan, a civil engineer at BUET, who is helping lead a Dutch-funded study of the Kobadak River project. The polder's land rose unevenly, with the area near the canal benefiting the most. Engineers are still deciphering sediment dynamics and considering how best to direct flooding, Khan says.

THE GREATEST CHALLENGE to opening the polders isn't the engineering, it's the people. Can some be persuaded to accept flooding on their land for years so that others can live flood-free?

Winners and losers are scattered around Beel Pakhimara. While farmers in nearby towns gather the rice harvest in golden fields now free from waterlogging, the beel is a desolate land of gray and pale green. Those who once farmed it have seen their rice paddies, fish farms, and ponds turn into a dumping ground for river sediment. It will remain that way until the canal is closed in approximately 2020, when the beel has filled with sediment and can be returned to cultivation.

Outside a small compound of brick and mud huts tucked against the embankment, 35-year-old Hanif Sardar (no relation to Jaharul) has little good to say about the project that's supposed to help rescue his country's land. He tried to get paid for the fifth of a hectare of land in the beel where he once grew rice that fed his family, he explains. "The government official says, 'Leave the application, something will happen. Don't worry.' But nothing happened," he says bitterly. "All the big landowners got the money. But the small landowners are not getting compensation."

The waterlogging that left a third of a meter of water in his yard for months each year ended in 2016. But that's outweighed by the family's loss of farmland. "The disadvantage is more," says Hanif's uncle, Amzad Sardar, "and the advantage is few."

To understand why some citizens turn against TRM projects, Mahmuda Mutahara, a Bangladeshi who recently earned her Ph.D., has spent much of the past 5 years traversing the region's pothole-riddled roads, often on the back of a motorcycle. Questioning residents, government officials, and others, she found government agencies disconnected from locals, spawning distrust and anger that have derailed controlled flooding attempts, sometimes spectacularly.

Government water officials "are not so much interested to talk to the local people," to learn about their wishes and explain how, in the long run, they might benefit, says Mutahara, who lives in Dhaka and studied environmental science at Wageningen University in the Netherlands. "That is the problem." To illustrate her point, Mutahara shares a story that for her is emblematic of the tangled social forces that make such work fraught with tensions.

Over the past decade, a rotating series of TRM projects on the nearby Hari River have repeatedly gone awry. Flooding in one beel dragged on years longer than planned. Farmers pressed for compensation for lost harvests. Fish-farming interests resisted the injection of fresh silt because it would fill their ponds. In 2012, a crowd rioted as government officials prepared to flood a beel there. The government suspended work and has yet to return.

Six years later, the Hari River problems remain, and so do the charred shells of three cars that rioters set on fire. "If, within the next 2 or 3 years, they do not do TRM upstream, the river will be silted up completely," Mutahara says.

The government has tried to learn from its mistakes, says Probir Kumar Goshwami, lead engineer at the Bangladesh Water Development Board's office in nearby Jessore. Last year, the government set up a temporary office near Beel Pakhimara where people could file for compensation. Water management officials are meeting more with locals. This year, the water agency plans to resume work on the beel at the center of the 2012 riot.

Despite some government resistance to TRM, Goshwami says there's no better alternative for those walled-off villages. "The tidal river management is the only solution."

Could controlled flooding be exported downstream to places like Polder 32? The chronic, low-level flooding on the Kobadak River differs from the threat of catastrophic, storm-induced flooding closer to the sea. And yet both places have something in common: a need to build up land.

As of now, no specific plan exists to bend to the water's will. Even a sort of "TRM lite"—a compartmentalized flooding in which small sections of polders are opened to raise elevation bit by bit—hasn't gained traction, says Haque, who has used computer models to study solutions to coastal vulnerability. Residents would have to be persuaded to live for years with flooded land. Reluctant government officials would need to embrace the approach. "This is a social problem," he says.

On Polder 32, the fields are dry today. It was one of the first places to get new, higher walls under the World Bank initiative. The breaches in the dike were closed in 2011.

Jaharul Sardar's view across the water, however, is a reminder that walls cannot erase risk. The land inside the polder is visibly lower than the shoreline outside it. But as he strolls barefoot atop the freshly rebuilt embankment, he feels safe for now. "If we have good embankments and cyclone shelters," he says, "we can survive." ■

This project was supported with a grant from the Pulitzer Center on Crisis Reporting.

NATURE'S STRATEGIES

A plant that stands and fights

Unlike those of us on legs, plants can't run away from what they don't like—yet they show remarkable resilience when under attack. Consider how the wild tobacco plant (*Nicotiana attenuata*, pictured above), a meter-high native of North America, protects itself from hungry insects. The plant senses the amino acid compounds in a caterpillar's saliva and responds with an alarm signal—a hydraulic or electrical pulse through its stems and leaves. Within minutes, the plant's cells rev up their production of nicotine, a poison that interferes with an animal's muscle function. When attacked, a single wild tobacco leaf can pack in a half a cigarette carton's worth of nicotine. But some caterpillars, such as hawkmoths, have evolved a way to pass that poison through their gut instead of absorbing it, forcing wild tobacco to unearth new countermeasures. The plant produces compounds that inhibit digestion and make the caterpillar sluggish, as well as abrasives that wear down the attacker's mouthparts. At the same time, the plant calls in help by emitting a scent that attracts ground-dwelling bugs and other caterpillar eaters, and then puts up chemical signposts to guide those predators to their already sluggish prey. Finally, a plant under siege redirects its resources, putting off flowering and growth until the caterpillars are gone. Amazingly, all of this is orchestrated not by a centralized brain, but by decision-making cells scattered throughout the plant.
—Elizabeth Pennisi

A dead coral skeleton is colonized by sponges and coralline algae. How can degraded coral ecosystems recover?

PERSPECTIVES

ECOLOGY

Seeking resilience in marine ecosystems

With recovery windows closing, how can reef corals resist climate change?

By **Emily S. Darling**[1,2] *and* **Isabelle M. Côté**[3]

Resilience is a popular narrative for conservation and provides an opportunity to communicate optimism that ecosystems can recover and rebound from disturbances. A resilience lens also reinforces the need for continued conservation investments, even in degraded ecosystems. It is probably for these reasons that resilience has become a conceptual cornerstone in the management of tropical coral reefs, which are one of the ecosystems most vulnerable to climate change.

The term "resilience" captures two dynamic processes: the ability of ecosystems to resist and absorb disturbance, and their ability to recover. Recent observations suggest that coral reef recovery is increasingly unlikely. The time windows for reefs to recover from consecutive mass bleaching events have shrunk from 25 to 30 years a few decades ago to just 6 years—far shorter than the 10 to 15 years that even the fastest-growing corals need to bounce back from catastrophic mortality (*1*). There has been some recovery from recent bleaching events on reefs that are isolated (*2*) or deeper and more structurally complex (*3*). Understanding more

[1]*Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, M5S 3B2, Canada.* [2]*Marine Program, Wildlife Conservation Society, 2300 Southern Boulevard, Bronx, NY 10460, USA.* [3]*Department of Biological Sciences, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada. Email: edarling@wcs.org*

PHOTO: EMILY DARLING/WCS

about how to catalyze such recovery processes is important. However, if climate change continues unabated, resilience may come not from the ability of coral reefs to recover but from their ability to resist.

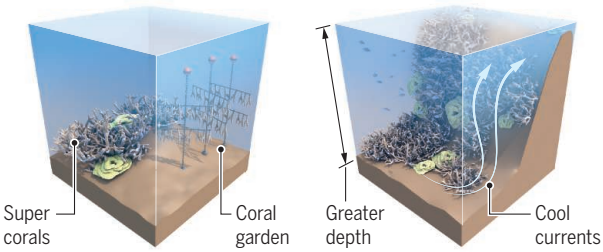
The idea of protecting resistant species or resistant areas is not new, but for various reasons, it is not often put into practice. To increase ecosystem resistance to climate change, there are two options: to increase intrinsic resistance, provided by traits that allow species to cope with a changing climate, or to increase extrinsic resistance, provided by locations that are less vulnerable to climate disturbances.

Coral resilience

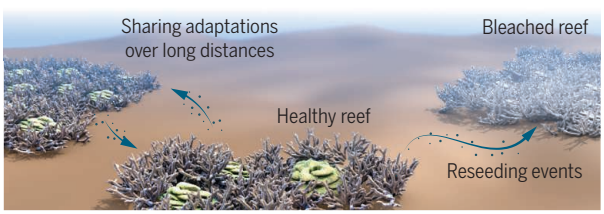
With shrinking recovery windows between bleaching events, future resilience may come not from the ability of corals to recover, but from their ability to resist the impacts of climate change.

Intrinsic resistance
Natural "super coral" traits could underpin new technologies.

Extrinsic resistance
Deeper reefs and colder waters might provide refuges for corals.



Connecting survivors
Even a few small, coral-rich refuges could act as sources of adaptations and propagules that can seed other reefs.



GRAPHIC: C. BICKEL/*SCIENCE*

INTRINSIC RESISTANCE: "SUPER" CORALS

Individuals or species that survive extreme climate events can have traits that underpin a general ability to cope or adapt to new environmental conditions. For corals, resistant traits include tolerance to warmer and acidified waters, salinity fluctuations, herbicides, diseases, and storms (*4*). These traits might be associated with aspects of the coral microbiome (*5*), or the ability of corals to draw on energy reserves or flexible feeding strategies (*6*). Natural populations of "super corals" that are tolerant of stressful conditions might arise after repeated bleaching events (*7*) or in more variable thermal environments, such as reefs exposed to tidal heat pulses (*8*).

A better understanding of natural super corals might lead to innovative technologies of assisted gene flow and assisted evolution that can help to climate-harden corals (see the figure). However, artificially increasing the resistance of corals to warmer and more acidic oceans through genetic engineering will be costly and controversial (*9*). All of these approaches are underway in the lab for a handful of the 1000 or so known species of corals.

Can such emerging technologies be used to cost-effectively restore high-diversity reefs at scales of hundreds or thousands of square kilometers? The answer right now is no, but efforts to assess the feasibility of large-scale restoration might change this answer in the future. A prime example is Australia's recent announcement of a 10-year Reef Restoration and Adaptation Program. As the options to protect corals from climate change vanish, it would be unwise to turn our backs on potentially beneficial technologies.

We can also help nature "pick the winners" by preserving the raw material for natural selection, i.e., genetic diversity. This can be done using a "portfolio-like" approach to conservation that maximizes diversity and connectivity among populations. Protecting populations across the broadest range of environments (including degraded reefs) can provide new combinations of genotypes or species, some of which might cope with future climate conditions (*10*).

EXTRINSIC RESISTANCE: A FEW GOOD REFUGES

Cool currents and deeper reefs might be areas where corals can persist. Today, however, such spatial refuges are increasingly rare. A case in point is the Great Barrier Reef, where successive bleaching events in 1998, 2002, and 2016 and 2017 have killed corals across vast areas, with no evidence that local management led to better outcomes (*11*). But some refuges still exist. Because of their locations, about 100, or 3%, of the 3800 reefs of the Great Barrier Reef system appear to have a relatively low risk of exposure to coral bleaching and predator outbreaks. These refuges are highly connected to other reefs by ocean currents. In a single reproductive event, these reefs could resupply larvae to up to half of all other reefs in the system (*12*).

Pessimists will caution against the precariousness of such a small number of reefs, but optimists will see "reefs of hope." Strategically protecting these few source reefs (see the figure) will not address the root cause of the coral reef crisis—greenhouse gas emissions—but it is a chance to support systematic resilience and avoid cascading failures across reef systems.

On a global scale, are a few refuges enough? The fossil record shows repeatedly that after historical mass extinction events, small, isolated, and ephemeral refuges protected critical remnant populations that seeded recovery and were crucial for the persistence of biodiversity (*13*). On contemporary coral reefs, even small areas of high coral cover can produce abundant, high-quality larvae (*14*). Finding and protecting even the smallest of resistant refuges is an urgent priority for global conservation efforts.

WHEN RESILIENCE RUNS OUT

Whether marine ecosystems resist, recover, restructure, or vanish hinges on how extreme future climate change is. It is almost certain that most of today's coral reefs will be transformed beyond recognition in the coming decades. Shifts in coral communities toward smaller, weedier species and dominance by other groups, such as algae and sponges, will alter ecosystem functioning and reduce the services that coral reefs provide to society (*15*).

These ecological shifts will force millions of people to adapt and change how they use and depend on the fisheries, tourism, and coastal protection provided by coral reefs. The political will necessary to improve the resilience of coral reefs or marine ecosystems might or might not materialize in time. Regardless, any fight for the remaining "reefs of hope" can, and must, occur alongside improving the resilience of people and communities to help dampen the coming climate shocks. ■

REFERENCES

1. T. P. Hughes *et al.*, *Science* **359**, 80 (2018).
2. J. P. Gilmour, L. D. Smith, A. J. Heyward, A. H. Baird, M. S. Pratchett, *Science* **340**, 69 (2013).
3. N. A. J. Graham, S. Jennings, M. A. MacNeil, D. Mouillot, S. K. Wilson, *Nature* **518**, 94 (2015).
4. M. J. H. van Oppen *et al.*, *Glob. Change Biol.* **23**, 3437 (2017).
5. T. D. Ainsworth, R. D. Gates, *Science* **352**, 1518 (2016).
6. A. G. Grottoli, D. Tchernov, G. Winters, *Front. Mar. Sci.* **4**, 215 (2017).
7. T. R. McClanahan, *Mar. Ecol. Prog. Ser.* **570**, 71 (2017).
8. L. J. Ruiz-Jones, S. R. Palumbi, *Sci. Adv.* **3**, e1601298 (2017).
9. K. Anthony *et al.*, *Nat. Ecol. Evol.* **1**, 1420 (2017).
10. M. S. Webster *et al.*, *Trends Ecol. Evol.* **32**, 167 (2017).
11. T. P. Hughes *et al.*, *Nature* **543**, 373 (2017).
12. K. Hock *et al.*, *PLOS Biol.* **15**, e2003355 (2017).
13. A. Godbold, S. Schoepfer, S. Shen, C. M. Henderson, *Geology* **45**, 607 (2017).
14. A. C. Hartmann, K. L. Marhaver, M. J. A. Vermeij, *Conserv. Lett.* 10.1111/conl.12410 (2017).
15. L. Alvarez-Filip, J. P. Carricart-Ganivet, G. Horta-Puga, R. Iglesias-Prieto, *Sci. Rep.* **3**, 3486 (2013).

10.1126/science.aas9852

ECOLOGY

What makes a terrestrial ecosystem resilient?

A complex set of biotic and abiotic factors determines the resilience of an ecosystem

By **Kathy J. Willis,[1,2] Elizabeth S. Jeffers,[2] Carolina Tovar[1]**

With increasing incidence of extreme climatic events, disease outbreaks, and other environmental perturbations, conservation of terrestrial ecosystems that can retain their structure and function despite environmental shocks has moved rapidly up the international political agenda. International environmental policies and targets such as the Aichi Biodiversity Targets and the Sustainable Development Goals include conserving resilient ecosystems as a key priority.

An ecosystem can display resilience in at least two ways: in the ability to resist an environmental perturbation and not switch to another state, and in how quickly it recovers after the disturbance (*1*). However, research into what makes a terrestrial ecosystem resilient is complex. Many hypotheses have been proposed, suggesting a suite of possible abiotic and biotic attributes responsible for resilience (*1*, *2*). Understanding resilience also requires comparative data sets that span both space and time and that address at least three questions: Where are the most resilient ecosystems, what attributes make them more resilient than others, and how close is an ecosystem to losing its resilience?

The first step toward addressing these questions is to find data sets demonstrating that particular ecosystems display resilience. Some projects fail at this hurdle. It is often difficult to find data sets spanning a long enough time to record the response of an ecosystem to an environmental perturbation, especially in forested ecosystems with trees that have decades-long generation times. Furthermore, the record of a resistant ecosystem will display no change despite a known climatic perturbation and is thus often discarded; a data record that shows no change is harder to publish.

In recent years, new sources of ecological data and algorithms for analyzing resilience have begun to provide answers to the three questions posed above. These studies have been carried out in almost every terrestrial biome (*3*); here, we focus on those undertaken in tropical ecosystems.

WHERE ARE THE MOST RESILIENT TROPICAL ECOSYSTEMS?

Data sets normally used to examine other ecological phenomena have recently been analyzed to determine both forms of resilience. Cole *et al.* (*4*), for example, used data from fossil pollen records to compare tropical forest recovery rates after perturbations between and within regions. In a meta-analysis, they examined 283 forest disturbance and recovery events over the past 20,000 years in Central and South American, African, and Asian rain forest blocks. They found substantial spatial differences in recovery rates; Central American tropical rain forests appeared to recover faster from environmental perturbations than those in South America and Asia.

Species richness, as in this West African tropical rain forest, may not always provide resilience to external perturbations (see the figure).

PHOTO: FABIAN PLOCK/ALAMY STOCK PHOTO

In another study, Poorter *et al.* (*5*) examined recovery rates of 45 Neotropical forest sites after clearance from 1500 records spanning the past 20 years. Even on this shorter time scale, the authors found strong geographical and climatic variation in recovery rates, with seasonally dry forests appearing to have less resilience than that of humid tropical lowland forests.

Innovative analysis of satellite imagery is starting to provide another important data set for determining spatial patterns of resilience (*6*, *7*). For example, Seddon *et al.* (*6*) used monthly Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data, taken globally at 5-km resolution between 2000 and 2013, to develop a vegetation sensitivity index that captures the relationship between climatic anomalies and relative variance of the vegetation (see the figure, left). This method identifies how sensitive different regions of vegetation have been to climatic variability over the past 14 years. In the tropics, more resilient areas included the woody savannas of the Brazilian Cerrado and the drylands of the Sahel. By contrast, vegetation in parts of the West Africa and the Amazon basins was highly sensitive to climatic perturbations.

WHAT ATTRIBUTES PROVIDE RESILIENCE?

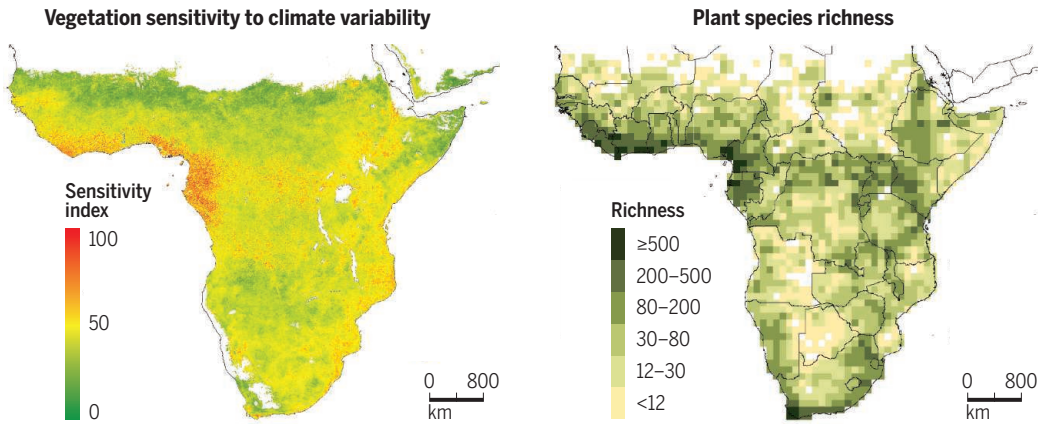
In addition to showing relative patterns of resilience, these spatial and temporal records provide important data sets with which to test the many hypotheses relating to resilience (*1*, *2*) and the complexity of the interacting biotic and abiotic factors that can lead to it (*2*). For example, a key abiotic attribute hypothesized to influence resilience is climate, and there is some evidence emerging to support this. In Neotropical dry forests, the highest rates of above-ground biomass recovery after clearance occurred in regions with higher local rainfall and lower water deficit (the difference between potential evapotranspiration and rainfall) (*5*).

Other studies have hypothesized that the number of times an ecosystem is disturbed, the less resilient it becomes, as indicated by a slowing down in recovery rates after each subsequent disturbance. However, this relationship does not always hold true, especially over longer time intervals. Cole *et al.* (*4*), for example, found the opposite effect in the fossil tropical forest data sets; the more times a system was disturbed, the faster it recovered, presumably because the vegetation became dominated by forest species that could tolerate and respond quickly to disturbance. The type of

[1]Royal Botanic Gardens, Kew, Richmond TW9 3AB, UK. [2]Oxford Long-term Ecology Laboratory, Department of Zoology, University of Oxford, Oxford OX1 3PS, UK. Email: k.willis@kew.org

No insurance policy, necessarily

Recent data suggest that at a continental scale, those regions in Africa rich in tropical plant species (*13*) are also the most sensitive to climate variability (*6*), implying that higher species richness does not necessarily lead to greater resilience.



GRAPHIC: J. YOU/*SCIENCE*

disturbance apparently can also influence the resilience of an ecosystem, and this varies according to vegetation type. For example, in tropical grasslands such as those in West Africa, disturbance by fires buffers ecosystems from forest encroachment and thus promotes grassland resilience (*8*). By contrast, frequent disturbance by fires in tropical forest-savanna transition zones can lead to loss of resilience in forest communities (*9*).

Another abiotic attribute hypothesized to account for ecosystem resilience is soil type. Again, there is some evidence to support this. In Poorter *et al.*'s study, high soil fertility had a positive influence on biomass recovery in Neotropical secondary forest plots (*5*). Similarly, a modeling study that incorporated remote sensing and field data predicts that rain forest situated on soils with low clay content will be least affected by an increase in the length of the dry season and will thus have higher resilience (*10*). Belowground biotic attributes may also be important in determining the resilience of an ecosystem; in particular, plants that have root systems associated with mycorrhizal fungi may have greater resilience to water stress in tropical dry forests (*11*).

There is also a suite of biotic factors to consider. Possibly the most widely cited is the insurance hypothesis. This suggests that more biodiverse ecosystems will be more resilient to environmental perturbations because they contain a greater number of species available to replace functions carried out by lost species. This certainly appears to be the case at the community level in some ecosystems (*12*) but does not necessarily hold at the continental scale. For example, regions with the highest tropical plant species richness in Africa (see the figure, right) (*13*) appear to be most sensitive to climate perturbations (see the figure, left) (*6*)—the opposite finding to the insurance hypothesis.

According to another biotic hypothesis, it is the characteristics of the component species (such as wood density, rooting depth, and leaf-area index) that make ecosystems more resilient (*1*). Here, some clear trends are starting to emerge. For example, Greenwood *et al.* (*14*) found that across forested biomes, mortality rates after drought were lower for species with greater wood density and lower specific leaf area. A global meta-analysis also identified these two characteristics as important for withstanding drought in tropical rain forests, whereas in tropical grasslands, plants with deeper roots were more resilient to drought (*3*).

HOW CLOSE IS A SYSTEM TO LOSING RESILIENCE?

Determining which biotic and abiotic factors contribute to resilient ecosystems is important for maintaining and enhancing them. However, when determining conservation strategies, it is also critically important to be able to identify when an ecosystem is about to lose its resilience and cross a threshold from a desirable to an undesirable stable state. Several methods have been proposed to do this. For example, at the continental scale, Hirota *et al.* (*15*) have shown that tropical and subtropical ecosystems in Africa, Australia, and South America switch to a savanna state when forest cover is less than 60%. This has direct implications for the management of tropical forests, where deforestation is a huge issue.

Another proposed method is to examine recovery rates from disturbances on the basis of the hypothesis that the closer a system is to a threshold, the slower the recovery rate will be (*1*). This approach appears to work in models but only seems to hold true for some terrestrial ecosystems. For example, Verbesselt *et al.* (*7*) found that recovery rates from perturbations slowed down sharply once mean annual precipitation fell below 1500 mm in evergreen tropical forests in South America, Africa, and Southeast Asia, possibly indicating a tipping point about to be crossed. By contrast, another study in the Amazon rain forests found only evidence of gradual change to several transitional forest states in response to a lengthening of the dry season (*10*), rather than an abrupt change from forest to another state.

There is also the fundamental question of whether switching between alternative states is always necessarily a bad thing. Recent studies indicate that tropical grasslands persist in a permanent transitional state and that the ability to switch between forest and savanna in response to perturbations underpins their resilience (*9*).

The recent studies discussed above are starting to test the many hypotheses that exist to explain resilience in terrestrial ecosystems. They reaffirm the complexity of resilience but also provide clear pointers for future research and conservation. In tropical ecosystems, soil type, belowground processes, and rooting depth are potentially important areas of future research with direct management applications. The factors responsible for resilience of tropical grasslands are another knowledge gap needing more research. Given the importance of terrestrial ecosystem resilience to natural resource security and supply across the globe, research into the attributes underpinning it should be high on any international agenda. ■

REFERENCES

1. D. Hodgson, J. L. McDonald, D. J. Hosken, *Trends Ecol. Evol.* **30**, 503 (2015).
2. B. L. Timpane-Padgham, T. Beechie, T. Klinger, *PLOS ONE* **12**, e0173812 (2017).
3. K. J. Willis *et al.*, in *State of the World's Plants, 2017 Report*, K. J. Willis, Ed. (Royal Botanic Gardens, Kew, 2017), pp. 42–49.
4. L. E. S. Cole, S. A. Bhagwat, K. J. Willis, *Nat. Commun.* **5**, 3906 (2014).
5. L. Poorter *et al.*, *Nature* **530**, 211 (2016).
6. A. W. R. Seddon, M. Macias-Fauria, P. R. Long, D. Benz, K. J. Willis, *Nature* **531**, 229 (2016).
7. J. Verbesselt *et al.*, *Nat. Clim. Change* **6**, 1028 (2016).
8. P. Laris, S. Dadashi, A. Jo, S. Wechsler, *Plant Ecol.* **217**, 583 (2016).
9. I. Oliveras, Y. Malhi, *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150308 (2016).
10. N. M. Levine *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 793 (2016).
11. K. Allen *et al.*, *Environ. Res. Lett.* **12**, 023001 (2017).
12. M. Hisano, E. B. Searle, H. Y. H. Chen, *Biol. Rev.* **93**, 439 (2018).
13. H. P. Linder, *Front. Ecol. Evol.* 10.3389/fevo.2014.00038 (2014).
14. S. Greenwood *et al.*, *Ecol. Lett.* **20**, 539 (2017).
15. M. Hirota, M. Holmgren, E. H. Van Nes, M. Scheffer, *Science* **334**, 232 (2011).

10.1126/science.aar5439

IMMUNOTHERAPY

Engineering a designer immunotherapy

Tailoring cytokine selectivity by engineering receptor-ligand pairs circumvents toxicity

By **Crystal L. Mackall**

Understanding of the human immune system has fueled development of numerous drugs, categorized as biologics, designed to mimic or inhibit natural immune responses. Recently, immunoengineering has emerged as a distinct discipline aimed at creating increasingly sophisticated therapeutics that can tailor immune responses for increased potency and/or diminished toxicity. On page 1037 of this issue, Sockolosky *et al.* (*1*) test the hypothesis that an engineered interleukin-2 (IL-2)–IL-2 receptor-β (IL-2Rβ) pair can expand effector T cells for cancer immunotherapy while avoiding the toxicities associated with administration of natural IL-2.

Immunoengineering has demonstrated remarkable clinical success. In 2017, the U.S. Food and Drug Administration (FDA) approved blinatumomab, a bispecific antibody that binds CD19, which is expressed on acute lymphoblastic leukemia (ALL) cells, and binds CD3 on resident, but antigen-nonspecific, T cells, resulting in T cell killing of ALL cells and improved clinical outcomes (*2*). Additionally, the FDA approved tisagenlecleucel and axicabtagene ciloleucel, engineered T cells expressing chimeric antigen receptors that merge monoclonal antibody binding domains with T cell signaling endodomains to induce potent antitumor effects in tumors expressing the CD19 antigen (*3*–*5*). Both of these agents mediate impressive clinical activity but are associated with considerable toxicities related to excessive cytokine production (*6*).

Recombinant human IL-2 (rhIL-2) is an FDA-approved biologic for the treatment of renal cell carcinoma and malignant melanoma. However, its clinical impact has been limited because the doses of rhIL-2 required for clinically meaningful antitumor effects induce profound capillary leakage syndrome, presumably related to IL-2 signaling in endothelial cells (*7*, *8*). Furthermore, although rhIL-2 induces the expansion of activated effector T cells and natural killer (NK) cells, both desirable for antitumor immunity, the agent also expands regulatory T cells (T_{regs}) (*9*), which inhibit antitumor immunity.

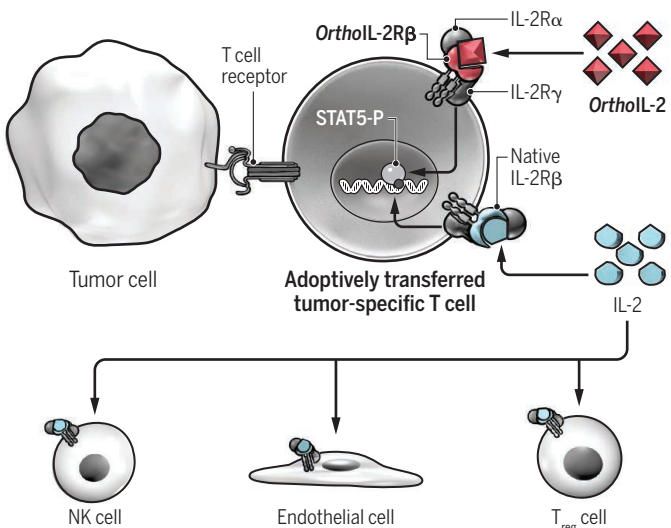
Sockolosky *et al.* sought to harness the benefits of IL-2 for cancer immunotherapy, while avoiding the toxicity and undesirable expansion of T_{regs}, by engineering an orthogonal IL-2 and IL-2Rβ pair. Orthogonality, a term originally used in mathematics to describe perpendicular vectors that are therefore nonintersecting, has evolved to also generally describe relationships that are statistically independent, partitioned, or isolated. Thus, orthogonal receptor–ligand pairs are restricted to exclusively signal together. Working in a mouse model and guided by structural knowledge of IL-2–IL-2Rβ binding, the authors created *ortho*IL-2Rβ, a mutant with two amino acid substitutions rendering it unable to bind natural IL-2. This property was independent of IL-2Rα, a nonsignaling component of the IL-2R complex that binds IL-2 with high affinity. *Ortho*IL-2Rβ retained normal interactions with IL-2Rγ, the subunit essential for transmitting the IL-2 signal via phosphorylated signal transducer and activator of transcription 5 (STAT5) (see the figure). The authors also randomly mutagenized IL-2 and selected two *ortho*IL-2s that selectively bound *ortho*IL-2Rβ. 3A10 was truly orthogonal as it showed no binding to natural IL-2Rβ yet bound *ortho*IL-2Rβ, albeit with reduced affinity compared to native IL-2–IL-2Rβ interactions. Conversely, 1G12 showed high affinity for *ortho*IL-2Rβ but also bound native IL-2Rβ with low affinity.

In a model of adoptive T cell immunotherapy for melanoma, they administered *ortho*IL-2 to mice that received infusions of transgenic T cells that recognize a peptide expressed by the melanoma cells. Prior to infusion, the transgenic T cells were genetically engineered to express *ortho*IL-2Rβ, but they also retained expression of the natural IL-2R. Both *ortho*IL-2s selectively expanded *ortho*IL-2Rβ–expressing transgenic T cells in vivo, but the magnitude of the effect was greater with 3A10, despite low-affinity binding, presumably due to the lack of competition by cells expressing natural IL-2Rβ. Similarly, 3A10 administration was nontoxic, whereas 1G12 induced the generalized toxicity previously described with rhIL-2 therapy. Importantly, both *ortho*IL-2s induced antitumor effects in the mice that were equivalent to those of native IL-2.

These results have the potential for broad clinical application because they provide convincing evidence that an orthogonal cytokine system can be designed and implemented for therapeutic effect in cancer immunotherapy. The results pave the way for testing adoptive cell therapy using tumor-specific T cells genetically engineered to express *ortho*IL-2Rβ. The system could also eliminate the need for lymphodepletion in patients receiving adoptive cell therapy for cancer. Numerous studies have demonstrated that adoptively transferred T cells do not substantially expand in vivo unless patients are first treated

*Ortho*IL-2 signaling expands tumor-reactive T cells

Sockolosky *et al.* generated *ortho*IL-2 to selectively bind *ortho*IL-2Rβ and induce STAT5 phosphorylation (P). Adoptive transfer of tumor-specific T cells genetically engineered to express *ortho*IL-2Rβ, followed by treatment with *ortho*IL-2, resulted in selective expansion of tumor-specific T cells in vivo without affecting NK cells and T_{regs} and without IL-2–associated toxicity from signaling on endothelial cells.



Department of Pediatrics and Medicine, and Stanford Cancer Institute, 265 Campus Way, G3141A, Stanford University, Stanford, CA, USA. Email: cmackall@stanford.edu

GRAPHIC: A. KITTERMAN/*SCIENCE*

with lymphodepleting agents (*10*–*12*), which induce homeostatic cytokines such as IL-7 and IL-15 that drive cell expansion. Standard lymphodepleting regimens induce substantial short- and long-term toxicity, and immune reconstitution following such regimens is prolonged and incomplete in many patients (*13*). The results presented here raise the prospect that selective expansion of tumor-specific T cells engineered to express *ortho*IL-2Rβ using *ortho*IL-2 could eliminate the requirement for lymphodepletion.

The ability of the *ortho*IL-2 system to separate the effects of IL-2 on immune effectors (activated T cells or NK cells) from effects on T_{regs} could also yield considerable clinical value. By engineering *ortho*IL-2Rβ expression on T_{regs}, but not immune effectors, one could selectively expand adoptively transferred T_{regs} for the treatment of autoimmune disease. In support of this, Sockolosky *et al.* demonstrated that *ortho*IL-2 can selectively expand *ortho*IL-2Rβ–expressing T_{regs} ex vivo. The results also open the door for generating orthogonal pairs for other cytokine-cytokine receptors, such as IL-7–IL-7R, where one could potentially harness the desirable capacity of IL-7 to expand T cells while preventing undesirable, potentially oncogenic, signaling on immature B cells (*14*).

Despite the promise of this elegant system, several issues remain. Administration of non-native proteins poses a potential risk for immunogenicity, as has occurred with recombinant proteins that mimic natural biologics (*15*). Furthermore, although IL-2 signaling in adoptively transferred T cells is considered necessary and sufficient for antitumor reactivity, it remains possible that IL-2 signaling in accessory cells, such as NK cells, could contribute to the efficacy of rhIL-2 therapy and thus will be excluded by orthogonal signaling. It is also possible that constitutive, high-level *ortho*IL-2Rβ expression, driven by a viral promoter such as that used here, might induce untoward regulatory circuits. Such issues will no doubt be investigated in depth to further understand immune biology and to more precisely tailor immune responses for greater therapeutic benefits. ■

REFERENCES

1. J. T. Sockolosky *et al.*, *Science* **359**, 1037 (2018).
2. H. Kantarjian *et al.*, *N. Engl. J. Med.* **376**, 836 (2017).
3. S. S. Neelapu *et al.*, *N. Engl. J. Med.* **377**, 2531 (2017).
4. S. L. Maude *et al.*, *N. Engl. J. Med.* **378**, 439 (2018).
5. S. L. Maude *et al.*, *N. Engl. J. Med.* **371**, 1507 (2014).
6. D. W. Lee *et al.*, *Blood* **124**, 188 (2014).
7. R. L. White Jr. *et al.*, *Cancer* **74**, 3212 (1994).
8. C. Krieg *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11906 (2010).
9. H. Zhang *et al.*, *Nat. Med.* **11**, 1238 (2005).
10. R. J. Brentjens *et al.*, *Blood* **118**, 4817 (2011).
11. Y. Cui *et al.*, *Blood* **114**, 3831 (2009).
12. F. L. Locke *et al.*, *Mol. Ther.* **25**, 285 (2017).
13. F. T. Hakim *et al.*, *J. Clin. Invest.* **115**, 930 (2005).
14. C. Shochat *et al.*, *J. Exp. Med.* **208**, 901 (2011).
15. N. Casadevall, *N. Engl. J. Med.* **346**, 469 (2002).

10.1126/science.aas9434

BIOPHYSICS

*p*Hirst sour taste channels *p*Hound?

Newly discovered acid channels may finally resolve the sour taste reception mystery

By **Craig Montell**

Do you enjoy the tartness of apples or grapes? A touch of sour can be delicious. But, highly sour foods are repulsive. This reaction warns against consuming foods spoiled by bacterial growth, such as rancid milk. So, it is no surprise that many animals have a sense of sour taste. In humans, it is one of five basic tastes, which also include sweet, bitter, salty, and umami (the savory taste induced by L-glutamate). Many receptors and ion channels in taste buds that are critical for detecting these chemicals in foods are now known (*1*). However, a mammalian sour taste receptor has been elusive. On page 1047 of this issue, Tu *et al.* (*2*) reveal a previously unrecognized H^+-selective channel that functions in mouse taste receptor cells (TRCs, which occur in taste buds of the tongue) that are essential for sour taste. This protein, Otopetrin1 (OTOP1), was originally identified because of its requirement in the vestibular system to maintain balance and to perceive gravity and limb orientation (*3*). The work by Tu *et al.* not only provides a strong candidate for the mammalian sour taste receptor but also raises questions concerning the broader roles of OTOP channels.

"Although the preponderance of evidence supports that OTOP1 is the sour taste receptor, behavioral analysis is needed to be certain."

High concentrations of Na^+ and H^+ are perceived as salty and sour, respectively, and are sensed by type III TRCs in our taste buds through mechanisms different from sugars, bitter compounds, and L-glutamate, which are sensed by type II TRCs. These latter substances bind to and activate G protein–coupled receptors, which initiate signaling cascades that culminate with the TRPM5 (transient receptor potential cation channel subfamily M member 5) channel (*1*). By contrast, Na^+ and H^+ are detected by cation channels. In mammals, the low-salt sensor is an epithelial Na^+ channel (ENaC) (*4*). However, the H^+ channel has been enigmatic. In the mouse, the H^+-sensing TRCs express a distant cousin of TRPM5, called PKD2L1 (polycystic kidney disease 2–like 1 protein) (*5*) (see the figure). Type III TRCs respond to acids by activating a H^+ channel that is blocked by Zn^{2+}, a form of inhibition that is common to the only other known eukaryotic H^+ channel, hydrogen voltage-gated channel 1 (Hv1) (*6*, *7*).

Multiple sour taste receptors have been proposed, including PKD2L1, acid-sensing ion channels (ASICs), and hyperpolarization-activated cyclic nucleotide–gated (HCN) channels (*1*). One by one, the candidates have not stood up to experimental scrutiny. To solve the mystery of the sour-sensing channel, Tu *et al.* cataloged several dozen genes predicted to encode proteins with multiple transmembrane domains (indicative of a channel) that are expressed in PKD2L1-positive TRCs but not in TRCs expressing TRPM5. They introduced the candidate proteins into in vitro cell-expression systems, looking for a H^+-influx current induced by extracellular acidity and inhibited by Zn^{2+}. OTOP1—with 12 predicted transmembrane domains—had >200,000-fold selectivity for H^+ over Na^+ and, in contrast to Hv1, showed only minor voltage dependence (*2*). However, it is unresolved if OTOP1 is simply regulated (gated) by acid or whether there is greater regulatory complexity.

Is OTOP1 the long-sought-after sour taste receptor? Strongly favoring this conclusion, the acid-activated H^+ conductance specific to TRCs expressing the *Pkd2l1* gene is virtually eliminated in TRCs from *tilted* mutant mice, which have a missense mutation in the *Otop1* gene that impairs vestibular system function. Although the preponderance of evidence supports that OTOP1 is the sour

Department of Molecular, Cellular, and Developmental Biology, University of California, Santa Barbara, Santa Barbara, CA 93106-9625, USA. Email: cmontell@ucsb.edu

taste receptor, behavioral analysis is needed to be certain. Complicating the behavioral analyses is that high acid concentrations not only stimulate TRCs but also activate nociceptor (pain) afferents (*8*). Thus, in the absence of a repulsion to high sour taste, such as in the *tilted* mice, H^+ would still cause aversive behavior because it would activate nociceptive neurons through an OTOP1-independent mechanism.

Because low levels of sour are appealing, and high levels are aversive, do separate TRCs and H^+ channels respond to slightly and strongly acidic foods? This would be reminiscent of Na^+-responsive TRCs, whereby one type senses low Na^+ concentrations through the ENaC channel and stimulates feeding and another class senses high Na^+ concentrations and other salts through an unknown channel and induces repulsion (*4*, *9*). Alternatively, there may exist only one class of H^+-activated TRC. Slight tanginess would stimulate the TRCs slightly and promote feeding, whereas high acidity could induce a greater response of the same TRCs and cause repulsion. Type III TRCs are also required for sensing water and carbonation and contribute to detecting high salt concentrations (*9*–*11*). Are these the same, overlapping, or distinct PKD2L1-positive type III TRCs that express OTOP1? Does loss of OTOP1 affect any of these other tastes? Interestingly, acidification suppresses the water taste response (*10*).

The finding that OTOP1 is a H^+ channel raises questions about its role in the vestibular system. In the inner ear, the functions of the sacculus and utricle, which aid in the perception of linear accelerations and gravity, depend on otoconia. These extracellular calcium carbonate–containing crystals overlay the gelatinous otolithic membrane surrounding the stereocilia of hair cells. When the head tilts, movement of the otolithic membrane, weighed down by otoconia, deflects the stereocilia, depolarizing hair cells. Mutations that impair the H^+ conductance through OTOP1, which is expressed in support cells adjacent to the otolithic membrane, prevent formation of otoconia. OTOP1 is proposed to regulate intracellular Ca^{2+} concentrations (*12*) and might be required to attain high Ca^{2+} concentrations in globular substance vesicles, which are required for otoconia formation, in support cells. However, the relationship of OTOP1 and these vesicles to otoconia formation is unclear. Ca^{2+}-channel activity can be inhibited by intracellular acidification. Therefore, OTOP1 might contribute to a homeostatic mechanism to maintain pH. Opposite to functioning in H^+ influx, as in TRCs, it might promote H^+ efflux from support cells, to prevent intracellular acidification.

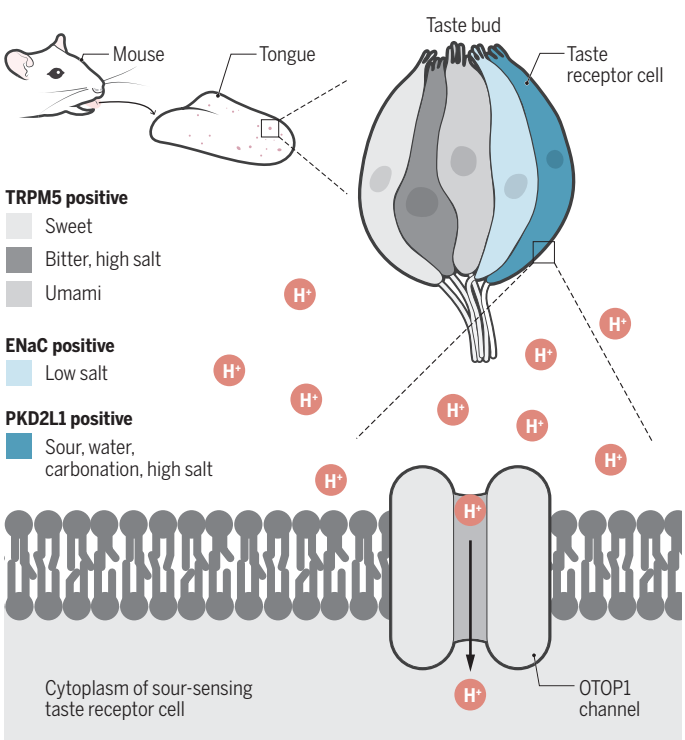
OTOPs comprise a conserved protein family, and Tu *et al.* found that mouse OTOP1 and OTOP2, human OTOP1, and a fly Otop each conduct H^+ upon lowering extracellular pH. OTOP family members are expressed in many cell types in humans and mice. Although intracellular acidification is typically cytotoxic, there are cells other than type III TRCs that may be endowed with H^+-influx channels. For example, osteoclasts, which are exposed to low extracellular pH (<5.5) in resorption pits in bones, have a H^+-influx pathway (*13*). However, it is only slightly inhibited by Zn^{2+} (*13*), differing from OTOP-dependent conductances.

A decrease in extracellular pH can accompany tissue injury or inflammation and enhance pain by increasing the activities of channels in nociceptors, such as TRPV1 (transient receptor potential cation channel subfamily V member 1), in peripheral sensory neurons. Because positive allosteric modulation of TRPV1 by mild acidification only occurs in response to extracellular acidification, OTOP-dependent H^+ influx might attenuate nociception. Indeed, OTOP1 expression is increased in dorsal root ganglion (DRG) neurons exposed to neuropathic conditions (*14*). However, H^+ influx would have to be moderate to prevent cytotoxic acidosis.

Given their wide expression patterns, loss of OTOPs might result in many pathologies. The intracellular pH of most cells is ~7.2, and the extracellular pH is usually ~7.4. However, cancer cells often have an increased intracellular pH (~7.4), whereas the extracellular pH is reduced (6.8 to 7.0) (*15*). This pH dysregulation can contribute to cancer by attenuating apoptosis and promoting cell proliferation and directed cell migration (*15*). Might loss of OTOPs contribute to certain cancers by limiting H^+ influx, which would otherwise reduce intracellular pH? Mouse OTOP2 and OTOP3 have somewhat distinct biophysical properties from OTOP1 (*2*). How does this relate to their biological roles? Do different OTOPs interact with distinct downstream effector proteins, which expand their functions? Clearly, there is much more to learn about this pHamily. ■

Sour taste reception

Mice and humans have five basic tastes: sweet, bitter, salty, umami, and sour. Many receptors and ion channels in taste buds are critical for detecting these chemicals in foods, but the sour taste receptor has been elusive. The H^+-selective channel OTOP1 is expressed in, and functions in, mouse acid-sensing taste receptor cells in taste buds of the tongue, and may be essential for sour taste.



GRAPHIC: N. DESAI/*SCIENCE*

REFERENCES

1. E. R. Liman *et al.*, *Neuron* **81**, 984 (2014).
2. Y.-H. Tu *et al.*, *Science* **359**, 1047 (2018).
3. B. Hurle *et al.*, *Hum. Mol. Genet.* **12**, 777 (2003).
4. J. Chandrashekar *et al.*, *Nature* **464**, 297 (2010).
5. A. L. Huang *et al.*, *Nature* **442**, 934 (2006).
6. M. Sasaki *et al.*, *Science* **312**, 589 (2006).
7. I. S. Ramsey *et al.*, *Nature* **440**, 1213 (2006).
8. T. Ohkuri *et al.*, *Chem. Senses* **37**, 523 (2012).
9. Y. Oka *et al.*, *Nature* **494**, 472 (2013).
10. D. Zocchi *et al.*, *Nat. Neurosci.* **20**, 927 (2017).
11. J. Chandrashekar *et al.*, *Science* **326**, 443 (2009).
12. E. Kim *et al.*, *J. Neurophysiol.* **104**, 3439 (2010).
13. M. Kuno *et al.*, *Pflugers Arch.* **468**, 837 (2016).
14. A. K. Reinhold *et al.*, *PLOS ONE* **10**, e0123342 (2015).
15. K. A. White *et al.*, *J. Cell Sci.* **130**, 663 (2017).

10.1126/science.aas9772

MICROBIOLOGY

Microbial warfare against viruses

Many new antiviral defense systems are found in bacteria and archaea

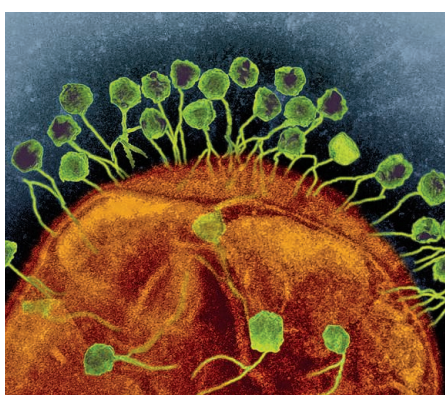
By **Jin-Soo Kim**[1,2]

Immune systems are divided into two types: innate and adaptive immunity. The innate immune system consists of components that are present in host organisms before infection by a pathogen and provides the first line of defense. For example, the restriction endonuclease system in prokaryotes degrades invader DNA (*1*). The adaptive immune system is activated after infection and provides immunological memory and long-lasting protection against the same pathogen. The CRISPR system captures a small piece of foreign DNA as a memory and recognizes and cleaves the same foreign DNA on reexposure (*2*), serving as adaptive immunity in prokaryotes. On page 1008 of this issue, Doron *et al.* (*3*) unveil previously unknown immune systems in bacteria and archaea against invading genetic elements, including plasmid DNA and bacteriophages (viruses that infect bacteria). Some of these systems may provide powerful tools for biomedical research and biotechnology, like restriction endonucleases and CRISPR.

Because genes in one immune defense system are often found next to genes in another defense system, forming "defense islands" in microbial genomes, Doron *et al.* hypothesized that unknown defense systems could be found near already-known antiviral systems. They computationally analyzed tens of thousands of genes in >45,000 bacterial and archaeal genomes and found dozens of candidate defense gene families with unknown functions in the vicinity of known defense systems. To characterize these gene families, they expressed candidate genes in two heterologous model bacteria, *Escherichia coli* and *Bacillus subtilis*. These recombinant bacteria were then challenged with a diverse set of bacteriophages. Remarkably, 9 of the 26 candidate systems conferred protection against at least one bacteriophage. Doron *et al.* named the nine gene families after legendary guardians such as Zorya and Thoeris, protective deities from Slavic and Egyptian mythologies, respectively.

These defense systems consist of either a single gene or multiple (up to 14) genes. Many of these encode proteins that recognize nucleic acids, such as helicases and nucleases, suggesting they have roles in recognition and destruction of foreign genetic elements. Interestingly, the Zorya system contains two genes encoding inner membrane proteins with flagellar motorlike domains. Doron *et al.* speculate that these two membrane proteins may form a proton (H^+) channel to depolarize membrane potential upon bacteriophage infection, leading to cell death—a previously unknown mechanism of antiviral defense in prokaryotes.

Thoeris is another interesting system that was identified by the enrichment of a gene family encoding the Toll-interleukin receptor (TIR) domain. The TIR domain is ubiquitously found in innate immune systems such as the Toll-like receptor and interleukin-1 receptor of invertebrates and mammals and many pathogen-resistance proteins in plants (*4*). These TIR proteins recognize diverse pathogen-associated molecular patterns (PAMPs) to trigger host immune responses. Doron *et al.* suggest that the Thoeris TIR protein is an evolutionary ancestor of these eukaryotic PAMP receptors. Unlike adaptive immune systems that are highly diverse, innate immune systems are conserved across the three kingdoms of life.



An electron microscopy image shows viruses (bacteriophage) surrounding a microbial cell.

PHOTO: AMI IMAGES/SCIENCE SOURCE

Different from the other nine antiviral systems, the Wadjet system, named after an Egyptian goddess, failed to provide *B. subtilis* protection against 10 bacteriophages tested by Doron *et al.*, but significantly reduced transformation efficiency of plasmid DNA, suggesting that it may target foreign plasmid DNA, a molecular parasite.

It is of particular interest to find out how these systems recognize foreign components, whatever they are, in a targeted manner. Some of these systems may find use as molecular tools, like the well-known members of the bacterial and archaeal immune systems. Historically, microbial defense systems have been successfully repurposed for broad applications in research, medicine, and biotechnology. For example, restriction endonucleases cut DNA precisely in vitro, enabling recombinant DNA technology (*5*). More recently, genome editing has been revolutionized by CRISPR–CRISPR-associated (Cas) systems (*6*–*9*). Perhaps the microbial defense genes found by Doron *et al.* will lead to innovations and applications in biotechnology and medicine.

There are likely to be other unknown defense genes hiding in microbial genomes. Doron *et al.* might have missed bona fide antiviral genes that are poorly expressed in the two model bacteria or that can protect other host bacteria against bacteriophages not tested in this study. The 10 validated defense systems are likely to be innate rather than adaptive immune systems, because adaptive immune systems like CRISPR cannot be identified by a single round of phage infection or plasmid transformation. The study of Doron *et al.* will encourage researchers to search for microbial defense genes and to investigate the molecular mechanisms of these defense systems. To do so, researchers will rely on molecular tools like CRISPR and on current and improved methods of reading, writing, and editing genes and genomes. To rephrase Sydney Brenner's remarks, with respect, progress in science is cyclic: new techniques leading to new discoveries, to new ideas, and then to new techniques again. ■

[1]*Center for Genome Engineering, Institute for Basic Science, Seoul, Republic of Korea.* [2]*Department of Chemistry, Seoul National University, Seoul, Republic of Korea.*
Email: jskim01@snu.ac.kr

REFERENCES AND NOTES

1. S. E. Luria, M. L. Human, *J. Bacteriol.* **64**, 557 (1952).
2. M. Jinek *et al.*, *Science* **337**, 816 (2012).
3. S. Doron *et al.*, *Science* **359**, eaar4120 (2018).
4. J. A. Hoffmann *et al.*, *Science* **284**, 1313 (1999).
5. S. N. Cohen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3240 (1973).
6. S. W. Cho *et al.*, *Nat. Biotechnol.* **31**, 230 (2013).
7. P. Mali *et al.*, *Science* **339**, 823 (2013).
8. L. Cong *et al.*, *Science* **339**, 819 (2013).
9. M. Jinek *et al.*, *eLife* **2**, e00471 (2013).

ACKNOWLEDGMENTS

J.-S.K. is supported by the Institute for Basic Science (IBS-R021-D1).

10.1126/science.aas9430

EARTH SCIENCE

Early plants and the rise of mud

Mudrock deposition in rivers increased by an order of magnitude after plants first evolved

By **Woodward W. Fischer**

The geological record of our planet provides evidence for a handful of ways in which life has fundamentally altered processes and environments at Earth's surface. It was the evolution of photosynthesis nearly 2.5 billion years ago that oxygenated the atmosphere and oceans (*1*), greatly increasing the spectrum of minerals found in rocks (*2*). Over the past 250 million years, the production of mineral skeletons by algae in the oceans transformed the way in which sediments accumulate in marine basins (*3*). On page 1022 of this issue, McMahon and Davies (*4*) illustrate how plants, too, have left an indelible mark in the geological record, their signature written in mud.

Mudrocks (fine-grained sedimentary rocks composed of silt- and clay-sized particles) are rare in the sedimentary deposits left by Precambrian and early Paleozoic (500 million years and older) rivers, in which sand and gravel are common (*5*); and the rise of mudrocks in river deposits in the geological record somehow reflects changes in the routing of sediment by rivers associated with the evolution of plants and their colonization of the landscape (*6*). This pattern formed part of the basis for interpretation of sedimentary deposits on Mars, created where ancient rivers drained into a lake at the bottom of Gale crater. The martian deposits—like those generated on Earth before plants—contain little in the way of mudrock (*7*). Although the sedimentary trend has been well-documented on Earth, it has remained poorly quantified. McMahon and Davies set out to measure the mudrock trend and narrow down its timing.

Geologists describe and quantify sedimentary rocks using measured stratigraphic sections, which are one-dimensional representations that capture the sequence and thickness of layers of sediment in the order in which they were deposited at a given location. McMahon and Davies collated stratigraphic sections from river deposits before and after plant evolution. They combined these observations with a survey of literature data to arrive at a data set covering hundreds of different sedimentary units, deposited on a range of continents over the past 3 billion years. From each of these sections, they extracted the percentage of mudrock present in the river deposits as a function of time.

The raw data readily recapitulated the expected pattern: Mudrocks were rare before the appearance of plants and common thereafter. But to obtain more quantitative information, statistical treatment of the data was required, because the number of stratigraphic units preserved in any given interval in the record is highly uneven. This type of bias is an omnipresent feature woven into the fabric of the geological record, and taking account of this is particularly important for terrestrial deposits, which are inherently rare (*8*). Using a couple of different approaches, McMahon and Davies demonstrate that the fractional abundance of mudrock rose by an order of magnitude or more following the evolution of plants. This reflects the tremendous impact that plants have had on the distribution of sediment in river corridors.

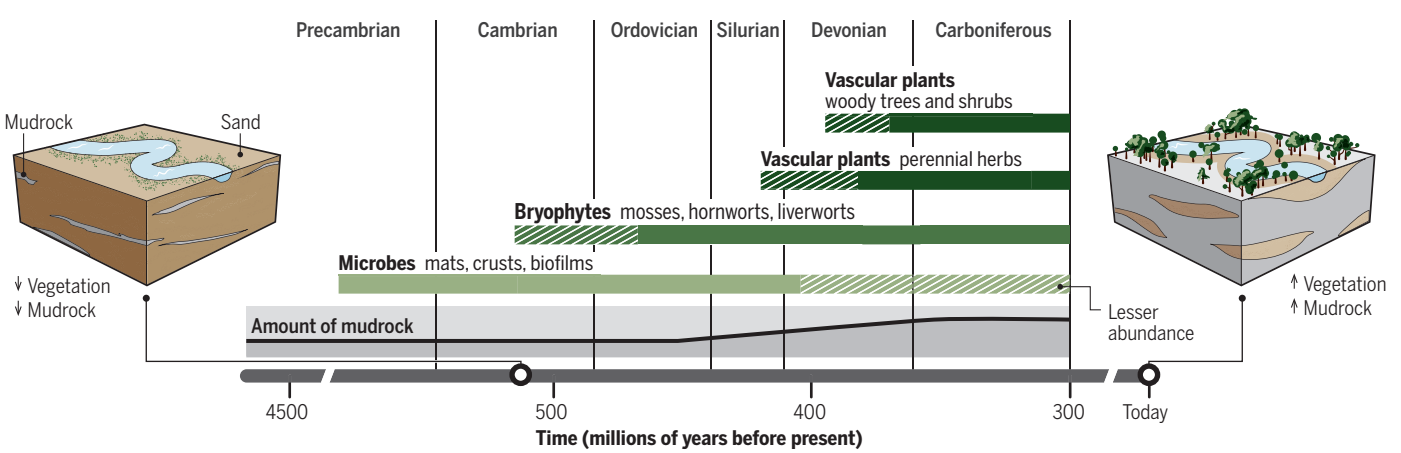
Particularly fascinating is the timing of this transition. From the estimates of McMahon and Davies, increase in mudrock began in the Late Ordovician to Silurian (450 to 420 million years ago). This implicates land plants, but is earlier than expected. The oldest fossil plants are Ordovician in age; earlier, more equivocal Cambrian-age microfossils hint at the earliest stages in plant evolution (see the figure). But it was not really until the Late Devonian (about 370 million years ago) that plant ecosystems were sufficiently well developed to be regarded as forests.

Early plant evolution occurred over a 100-million-year interval of increasing ecophysiological complexity and landscape occupation (*9, 10*). Primitive plants were mostly bryophytes (related to mosses and liverworts), and early vascular plants in the later Silurian and early Devonian (about 425 million years ago) had stem lengths measured in centimeters, little water-conducting tissue, no woody tissue, and were limited to wet environments. It is interesting that such primi-

Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA. Email: wfischer@caltech.edu

Plants reshaped the sedimentary deposits left by a river

Muddy floodplains were rare on prevegetated landscapes, compared with those developed after the evolution of plants and their colonization of the landscape. The rise of mudrock coincided with some of the earliest events in plant evolution.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

tive plant ecosystems would have substantial impact on the behavior of rivers and the construction of their floodplain deposits.

According to the results of McMahon and Davies, mudrock in river deposits continued to increase through the Devonian-Carboniferous interval (about 370 to 300 million years ago) during the evolution of deeply rooted plants and true forests. If the timing of the initial increase in mudrock is correct, it suggests that the earliest plants made a substantial impression on the landscape.

What was it about these early plant-bearing ecosystems that generated so much mudrock in their deposits? There are many ideas, and little is certain. McMahon and Davies point to processes by which plants may have increased mud production during erosion and weathering. With that in mind, there are preserved Precambrian soil horizons composed of mudrocks and plenty of Precambrian marine mudrocks indicating that substantial fluxes of mud passed from the continents to marine basins before plants evolved. Our planet, it seems, has always had enough mud to go around. It is therefore likely that early plants affected the mechanics of floodplain construction. For example, the presence of plants on the landscape decreases erosion rates; and thus it was long hypothesized that erosion—in particular by wind—removed sediment from prevegetated landscapes (*11, 12*). Even if mud was deposited on prevegetated floodplains, its removal by erosion might have been efficient.

In addition to inhibiting erosion, plants also interact with river flows and promote the deposition of fine-grained sediment. This can help armor riverbanks and slow their lateral migration; such a process might also aid in preserving muddy floodplain deposits.

Whatever the exact causes, the muddy signal that plants left in the geological record is prodigious. That such seemingly simple early-appearing plant ecosystems had such sedimentary impact means that there is still much to learn about the nature of the interactions and coevolution of terrestrial ecosystems and their underlying landscapes. ■

REFERENCES

1. W. W. Fischer, J. Hemp, J. E. Johnson, *Annu. Rev. Earth Planet. Sci.* **44**, 647 (2016).
2. R. M. Hazen, *Sci. Am.* **302**, 58 (2010).
3. A. H. Knoll, *Rev. Mineral. Geochem.* **54**, 329 (2003).
4. W. J. McMahon, N. S. Davies, *Science* **359**, 1022 (2018).
5. D. Winston, in *Fluvial Sedimentology–Memoir 5*, A. D. Miall, Ed. (Canadian Society of Petroleum Geologists, Calgary, 1977), p. 343.
6. M. R. Gibling, N. S. Davies, *Nat. Geosci.* **5**, 99 (2012).
7. J. P. Grotzinger *et al.*, *Science* **350**, aac7575 (2015).
8. S. E. Peters, *J. Geol.* **114**, 391 (2005).
9. P. Kenrick, *Science* **358**, 1538 (2017).
10. C. K. Boyce, J.-E. Lee, *Annu. Rev. Earth Planet. Sci.* **45**, 61 (2017).
11. S. A. Schumm, *Geol. Soc. Am. Bull.* **79**, 1573 (1968).
12. R. H. Dott Jr., *J. Geol.* **111**, 387 (2003).

10.1126/science.aas9886

OPTICS

Fermi arcs connect topological degeneracies

Surface or bulk Fermi arcs are engineered in photonic structures to connect ideal Weyl points or exceptional points

By **Şahin K. Özdemir**

In condensed matter and photonics, topology is defined with respect to the energy bands in momentum space (*1*). The boundary between two topologically different phases (for example, supporting right- versus left-handed particles) appears as degeneracies where two linear dispersion bands intersect. Fundamental point degeneracies in two-dimensional (2D) and 3D Hermitian systems—known as Dirac and Weyl points, respectively—have been observed in photonic structures (*2*–*6*) but not the ideal Weyl points and the helicoidal dispersion, which leads to the open Fermi arcs connecting points of opposite chirality. On page 1013 of this issue, Yang *et al.* (*7*) demonstrate an ideal Weyl system with four Weyl points (*8*, *9*) and helicoidal surface Fermi arcs interconnecting them in a 3D photonic crystal composed of metallic inclusions. On page 1009 of this issue, Zhou *et al.* (*10*) explore non-Hermitian topological photonics in which radiative losses come into play, and demonstrate the emergence of bulk Fermi arc and half topological charges in a 2D-periodic photonic crystal.

> ***"...the reported observations... will surely help to explore photonics in new regimes..."***

The observation of Weyl points at microwave (*2*, *3*) and optical frequencies (*4*) serves as a fingerprint of the topological nature of the corresponding photonic system and guarantees the existence of Fermi arcs, which act as pipelines connecting Weyl points of opposite chirality (*5*, *6*). However, further development and potential applications have been hindered by the lack of an ideal Weyl system (*8*, *9*) in which all Weyl points are symmetry-related, exist at the same energy, and are free from nontopological bands.

Despite both being degeneracies of a Hermitian system (that is, one with no energy exchange with the environment), Dirac and Weyl points differ in critical ways. For example, any perturbation that breaks parity (*P*) or time-reversal (*T*) symmetry lifts the 2D Dirac point degeneracy and opens a band gap; the Dirac point splits and loses its topological protection. However, creation of topologically protected Weyl points actually requires the breaking of *P*- or *T*-symmetry, or both (see the figure) (*11*, *12*). Hermitian perturbations only shift Weyl points but cannot lift their degeneracy. A Weyl point is annihilated only by bringing together two points of opposite chirality (for example, ones with chirality of +1 or –1).

The minimum number of Weyl points in systems respecting *P*-symmetry (system breaks *T*-symmetry) is two, whereas in systems respecting *T*-symmetry (system breaks *P*-symmetry), it is four. Weyl points with nonzero chirality have been previously reported in systems with broken *P*- and *T*-symmetry. However, Weyl points in those studies were not ideal because they were at different frequencies and were not symmetry-related.

Yang *et al.* succeeded to create four ideal Weyl points with different topological chirality (two with +1 and two with –1) in the microwave regime by using a specially designed meta-crystal that breaks *P*-symmetry but leaves *T*-symmetry untouched. They characterized the band structure of their meta-crystal using angle-resolved transmission spectroscopy by using a near-field antenna as the stationary excitation point-source on the crystal surface and measuring the near field with a scanning antenna on the opposite side. Fourier analysis of the acquired spatial distribution of the electric field at each angle and excitation frequency was used to determine the band structure of the surface states at each angle, confirming the linear band crossings of the surface states around the Weyl points.

In order to characterize the surface arcs, Yang *et al.* probed their meta-crystal using two different configurations of transmission near-field scanning system in which

Department of Engineering Science and Mechanics, The Pennsylvania State University, University Park, PA 16802, USA. Email: sko9@psu.edu

a near-field antenna scanned the top surface to map out the bulk and surface modes while an excitation source was placed either at the center of the bottom layer of the meta-crystal stack or at the edge of the top surface. These measurements revealed the presence of two helicoidal surface Fermi arcs connecting bulk states with opposite topological charges.

The concepts of topologically nontrivial phases have been developed for closed (Hermitian) physical systems, but in photonics, systems are open, and the perturbations can cause loss or gain of signal. Such non-Hermitian systems may have exceptional point (EP) degeneracies—where not only the eigenvalues but also the corresponding eigenvectors coalesce (*13*)—that are different from Hermitian degeneracies. Non-Hermiticity will likely provide a richer topological landscape with no counterpart in Hermitian systems (*14*).

Zhou *et al.* report the observation of a Fermi arc and its relation to the topological features of a non-Hermitian system. Different than the Fermi arc reported by Yang *et al.*, the Fermi arc in this study resides in the bulk dispersion of a 2D non-Hermitian system, rather than on the surface of a 3D Hermitian system, and originates from the non-Hermiticity of the system, rather than from the presence of Weyl points. Similar to the surface Fermi arc connecting Weyl points of opposite charges (±1), this bulk Fermi arc connects a pair of EPs with opposite half-charges (±1/2).

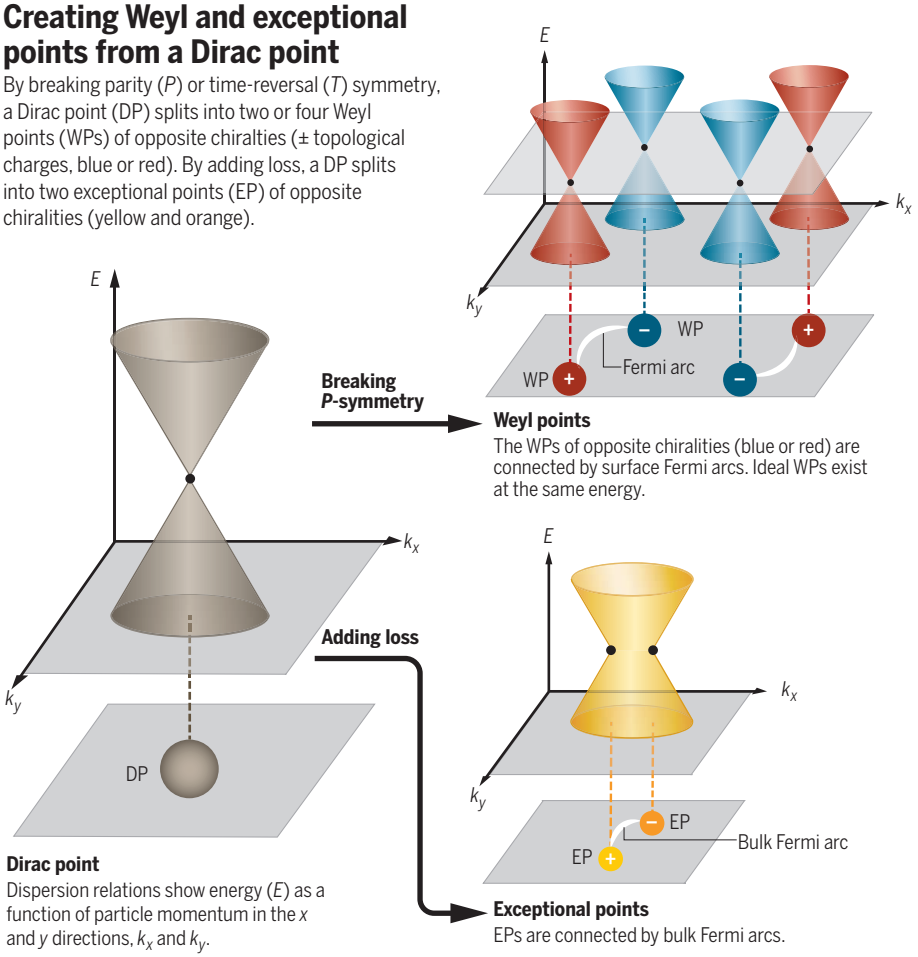
The EPs in the study of Zhou *et al.* were formed via radiative losses on a 2D periodic photonic crystal with a rhombic lattice having a Dirac point with nontrivial Berry phase. The losses transform the Dirac points into EPs. Zhou *et al.* characterized the sample using angle-resolved scattering measurements in which a tunable laser illuminated the sample and a charge-coupled device camera collected the scattered light. This direct probing revealed open-ended isofrequency contours connecting a pair of EPs, similar to surface Fermi arcs connecting the Weyl points. Far-field polarization measurements showed that the polarization vector experiences a 180° winding around the Fermi arc, corresponding to a half-integer topological charge that is a manifestation of the ±1/2 topological index of the EPs in the system.

The observations reported in this issue by Yang *et al.* and Zhou *et al.* are important advances in topological photonics for many reasons. First, the demonstration of ideal Weyl points and establishing the connection between topological concepts and non-Hermiticity open a broader avenue for fundamental research of the emergent fields of topological physics and photonics. Second, they hold the promise for practical applications. For example, ideal Weyl points may lead to enhancement of transport induced by chiral anomaly and are expected to be useful for invisibility cloaking and 3D-imaging, as well as provide better angular and frequency selectivity (*1*, *9*, *11*, *12*). Similarly, as suggested by Zhou *et al.*, half topological charges in polarization can be used for the generation of half-integer vector-vortex beams. Third, observation of a bulk Fermi arc and topological charge in an open system begins to establish and explore the connection between topological concepts and non-Hermitian systems. Fourth, the platforms developed will allow for the exploitation of distinctive properties of Weyl points in Hermitian systems and topological features in non-Hermitian systems to their full extent.

Although it remains to be seen how the reported observations will lead to novel devices and technological advances, they will surely help to explore photonics in new regimes—for example, by including nonlinearities and quantum effects—and lead to better ways of controlling and engineering light and its interaction with matter. There are still many issues to address. For instance, how can one realize ideal Weyl points in optical frequencies? Is there a connection between Weyl points and EPs? Last, how do non-Hermitian perturbations in the form of induced gain or loss affect Weyl points, topological charges, and surface Fermi arcs? ■

Creating Weyl and exceptional points from a Dirac point

By breaking parity (*P*) or time-reversal (*T*) symmetry, a Dirac point (DP) splits into two or four Weyl points (WPs) of opposite chiralties (± topological charges, blue or red). By adding loss, a DP splits into two exceptional points (EP) of opposite chiralities (yellow and orange).



Dirac point
Dispersion relations show energy (*E*) as a function of particle momentum in the *x* and *y* directions, k_x and k_y.

Weyl points
The WPs of opposite chiralities (blue or red) are connected by surface Fermi arcs. Ideal WPs exist at the same energy.

Exceptional points
EPs are connected by bulk Fermi arcs.

GRAPHIC: N. CARY/*SCIENCE*

REFERENCES AND NOTES

1. L. Lu, J. D. Joannopoulos, M. Soljačić, *Nat. Phys.* **12**, 626 (2016).
2. L. Lu *et al.*, *Science* **349**, 622 (2015).
3. B. Yang *et al.*, *Nat. Commun.* **8**, 97 (2017).
4. J. Noh *et al.*, *Nat. Phys.* **13**, 611 (2017).
5. P. J. W. Moll *et al.*, *Nature* **535**, 266 (2016).
6. X. Wan, A. M. Turner, A. Vishwanath, S. Y. Savrasov, *Phys. Rev. B* **83**, 205101 (2011).
7. B. Yang *et al.*, *Science* **359**, 1013 (2018).
8. J. Ruan *et al.*, *Nat. Commun.* **7**, 11136 (2016).
9. L. Wang, S.-K. Jian, H. Yao, *Phys. Rev. A* **93**, 061801(R) (2016).
10. H. Zhou *et al.*, *Science* **359**, 1009 (2018).
11. L. Lu, J. D. Joannopoulos, M. Soljačić, *Nat. Photon.* **8**, 821 (2014).
12. M. Z. Hasan, S.-Y. Xu, I. Belopolski, S.-M. Huang, *Annu. Rev. Condens. Matter Phys.* **8**, 289 (2017).
13. W. D. Heiss, *J. Phys. A Math. Gen.* **45**, 444016 (2012).
14. D. Leykam, K. Y. Bliokh, C. Huang, Y. D. Chong, F. Nori, *Phys. Rev. Lett.* **118**, 040401 (2017).

ACKNOWLEDGMENTS

S.K.O. is supported by ARO grant No. W911NF-18-1-0043 and Pennsylvania State University, Materials Research Institute.

10.1126/science.aar8210

POLICY FORUM

CLIMATE POLICY

Linking climate policies to advance global mitigation

Joining jurisdictions can increase efficiency of mitigation

By **Michael A. Mehling,[1,2] Gilbert E. Metcalf,[3,4,5] Robert N. Stavins[4,5,6]**

The November 2017 negotiations in Bonn, Germany, under the auspices of the United Nations Framework Convention on Climate Change (UNFCCC) validated that the Paris Agreement has met one of two necessary conditions for success. By achieving broad participation, including 195 countries, accounting for 99% of global greenhouse gas (GHG) emissions (*1*), the agreement dramatically improves on the 14% of global emissions associated with countries acting under the Kyoto Protocol (*2*), the international agreement it will replace in 2020. But the second necessary condition, adequate collective ambition of the nationally determined contributions (NDCs) that countries have individually pledged, has not been met. One promising approach to incentivize countries to increase ambition over time is to link different climate policies, such that emission reductions in one jurisdiction can be counted toward mitigation commitments of another jurisdiction. Drawing on our research and our experiences in Bonn, we explore options and challenges for facilitating such linkages in light of the considerable heterogeneity that is likely to characterize regional, national, and subnational policy efforts.

Linkage is important, in part, because it can reduce the costs of achieving a given emissions-reduction objective (*3*). Lower costs, in turn, may contribute politically to embracing more ambitious objectives. In a world where the marginal cost of abatement (that is, the cost to reduce an additional ton of emissions) varies widely, linkage improves overall cost-effectiveness by allowing jurisdictions to finance reductions in other jurisdictions with relatively lower costs while allowing the former jurisdictions to count the emission reductions toward targets set in their NDCs (see the figure). For example, the baseline efficiency of energy use in low-income countries is very low, relative to high-income countries. Linking can leverage such differences to reduce overall mitigation cost. In effect, linkage drives participating jurisdictions toward a common cost of carbon, equalizing the marginal cost of abatement and producing a more efficient distribution of abatement activities. These benefits could potentially reduce the cost of achieving the emissions reductions specified in the initial NDCs under the Paris Agreement 32% by 2030 and 54% by 2050 (*4*).

In addition to lowering costs, linkage can improve the functioning of individual markets, reducing market power by including more firms and reducing price volatility by enlarging the market. Beyond such direct economic benefits, political benefits exist. As jurisdictions band together, linking can signal political momentum that contributes to more policies where they do not yet exist and more ambitious policies where they are already in place. Also, administrative economies of scale can be achieved through knowledge sharing in policy design and operation and through shared administrative costs. Finally, and importantly, linkage can allow for the key UNFCCC equity principle of "common-but-differentiated responsibilities and respective capabilities" to be pursued without sacrificing cost-effectiveness.

There are also legitimate concerns with linkage, including distributional impacts within and across jurisdictions, even though aggregate abatement costs are reduced. Because linking is inherently voluntary, however, linking will generally not occur unless both parties to a link anticipate that overall benefits of the link, including revenue from selling emission reductions, will outweigh costs. Likewise, individual exchanges made between compliance entities are voluntary.

Transferring pollution obligations can raise concerns about environmental justice. Although GHGs are a global pollutant, changes in GHG emissions can affect emissions of correlated local pollutants (for example, particulate matter). This is a reasonable concern, but linkage could help reduce correlated local pollution in developing countries, because jurisdictions that take on increased mitigation efforts as a result of linkage, many of which will be low-income developing countries, will see local pollution decrease along with lower GHG emissions.

A more serious concern of linkage stems from the automatic propagation of some design elements from one system to another, in particular, cost-containment mechanisms in cap-and-trade systems—banking, borrowing, and price collars. This means that there is decreased autonomy, as rules in one system can affect prices in another. All of this refers to what we think of as "hard linkage," a formal recognition by a mitigation program in one jurisdiction of emission reductions undertaken in another jurisdiction for purposes of complying with the first jurisdiction's mitigation program. Examples of hard linkage are the links between cap-and-trade systems in the state of California, USA, and Québec, Canada, and, more recently, the European Union (EU) and Switzerland.

But another possibility is "soft linkage," by which we mean an agreement—explicit or implicit—to harmonize carbon prices either at a specific level or within overlapping bands. With soft linkage, there is no recognition of emission reductions in one system by the other system for purposes of compliance. Still, by aligning carbon prices, such harmonization improves overall economic efficiency.

LINKAGE IN THE PARIS AGREEMENT

Article 6.2 of the Paris Agreement provides a foundation for linkage by recognizing that parties to the agreement may "choose to pursue voluntary cooperation in the implementation of their" NDCs through "the use of internationally transferred mitigation outcomes" (ITMOs) (*5*). In contrast to the Kyoto Protocol (which also includes provisions for international cooperation), the voluntary and flexible architecture of the Paris Agreement allows for wide variation, not only in the types of climate policies countries choose to implement but also in the form and stringency of the abatement targets they adopt.

To be clear, there are three conceptually—and operationally—distinct aspects of international policy linkage: (i) (the focus of our analysis) provisions in Article 6.2 of the Paris Agreement and related guidance that can facilitate international linkage, by providing, for example, for ITMOs to be used as an accounting mechanism when "compliance" with NDCs is measured; (ii) agreements between two or more jurisdictions to recognize emission reductions generated in the other jurisdictions; and (iii) two or more compli-

[1]*Massachusetts Institute of Technology, Cambridge, MA, USA.* [2]*University of Strathclyde, Glasgow, UK.* [3]*Tufts University, Medford, MA, USA.* [4]*National Bureau of Economic Research, Cambridge, MA, USA.* [5]*Resources for the Future, Washington, DC, USA.* [6]*Harvard University, Cambridge, MA, USA. Email: robert_stavins@hks.harvard.edu*

ance entities, one in each of the linked jurisdictions, engage in an exchange, for example, permitting allowances to move between cap-and-trade systems.

Linkage is relatively straightforward when policies involved are similar. But there are several potential sources of heterogeneity: type of policy instrument (e.g., taxes, cap-and-trade, performance or technology standards); level of government jurisdiction involved (e.g., regional, national, subnational); status under the Paris Agreement (that is, whether or not the jurisdiction is a party to the agreement or within a party); nature of the policy target (e.g., absolute mass-based emissions, emissions intensity, change relative to business-as-usual); and operational details of the country's NDC (e.g., type of mitigation target, choice of target and reference years, sectors and GHGs covered). Most forms of heterogeneity, however, do not present insurmountable obstacles to linkage.

In principle, the most straightforward case of international climate policy linkage would be a pair of national cap-and-trade systems in parties to the agreement, with each using an absolute (mass-based) target in its NDC (for example, cap-and-trade systems in New Zealand and Switzerland). A less obvious case would be a pair of subnational policies—one a carbon tax and one a cap-and-trade system (for example, carbon tax in British Columbia, Canada, and cap-and-trade in Tokyo, Japan). Both policies can be designed to facilitate heterogeneous linkage (*6*). Another case of heterogeneous linkage might be between the EU emissions trading system and California's cap-and-trade program. All of these would be conceptually feasible and merit consideration, although each raises issues that require attention and call for specific accounting guidance, if linkage is to include the use of ITMOs under the Paris Agreement.

A PATH FORWARD

Parties to the Paris Agreement are working to elaborate guidance on Article 6.2 but have expressed widely differing views on what issues to include (*7*). In Bonn, parties signaled agreement on the need to offer at least minimal guidance on how to account for transfers of ITMOs. Beyond that, however, positions diverge on whether to address broader questions that bear on linkage under Article 6.2. Particular divisions center around issues of environmental integrity, governance, and sustainable development.

Our analysis, based on case studies of various types of heterogeneous linked systems, reveals common themes (*3*). What emerges is the importance of guidance on Article 6.2 that sets out a robust accounting framework to prevent double counting of GHG reductions, to ensure that the timing (vintage) of claimed reductions and respective ITMO transfers is correctly accounted for, and to ensure that participating countries make appropriate adjustments for emissions or reductions covered by their NDCs when using ITMOs. Suggested approaches for ITMO accounting under Article 6.2 (*3*) include, in particular, how to make adjustments to national emission budgets to account for ITMOs and how to account for heterogeneous base years, different vintages of targets and outcomes, and transfers between parties and non-parties to the agreement. These issues were identified, if not yet resolved, during the negotiations on Article 6.2 in Bonn. As negotiators proceed to address them, they can draw on a wealth of experience and existing research (*8*–*11*). Future study should expand on the specific conditions of accounting and ITMO transfers under the evolving architecture of the Paris Agreement.

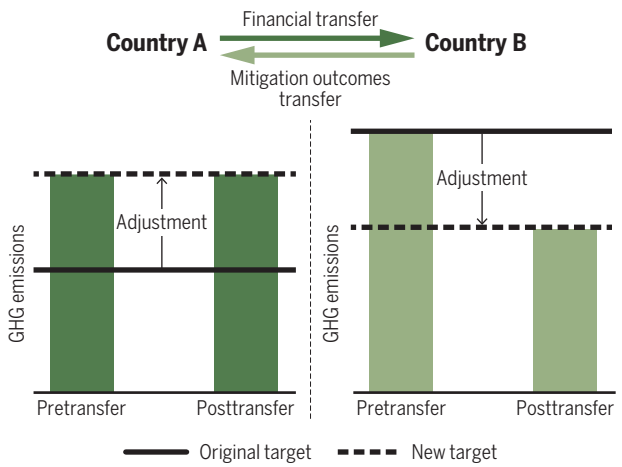
An important insight from our analysis, however, is that parties should exercise caution when developing guidance that goes beyond accounting issues. Onerous conditions related to the ambition or integrity of domestic action, for instance, could deter linkage. This does not mean that such concerns should be neglected; rather, they are best addressed under separate corresponding negotiating tracks—such as the Talanoa dialogue, to take stock of collective efforts of parties—or the enhanced transparency framework under Article 13 of the Paris Agreement.

Clear and consistent guidance for accounting of emissions transfers under Article 6.2 can contribute to greater certainty and predictability for parties engaged in voluntary cooperation, facilitating expanded use of linkage. Too much guidance, however, particularly if it includes restrictive quality or ambition requirements, might impede linkage and dampen incentives for cooperation. Such a combination of common accounting rules and an absence of restrictive criteria and conditions may accelerate linkage and allow for broader and more ambitious policy cooperation, which can increase the potential for parties to scale up the ambition of their NDCs. That may ultimately foster stronger engagement between parties (and non-parties), as well as with regional and subnational jurisdictions.

The parties to the Paris Agreement will continue negotiations in May, toward a goal of agreeing to a finalized rule book for Article 6 at the annual UNFCCC summit in Katowice, Poland, in December 2018. Decisions that the negotiators reach this year could greatly advance, or impede, international climate policy linkage and thereby play a key role in determining the fate of the Paris Agreement. ■

Transferring mitigation outcomes

Rather than reduce its greenhouse gas (GHG) emissions to meet its original target, country A cooperates with country B, which can reduce emissions at lower cost. Incentivized to reduce emissions, country B sells part of its mitigation outcome to country A. Both targets are adjusted to reflect the transfer, and country A meets its adjusted target at lower cost than if it had reduced its own emissions.



GRAPHIC: K. SUTLIFF/*SCIENCE*

REFERENCES AND NOTES

1. UNFCCC, "Paris Agreement—Status of Ratification"; http://unfccc.int/paris_agreement/items/9444.php [accessed 3 January 2018].
2. UNFCCC, "Kyoto Protocol"; http://unfccc.int/kyoto_protocol/items/2830.php [accessed 20 November 2017].
3. M. A. Mehling, G. E. Metcalf, R. N. Stavins, "Linking Heterogeneous Climate Policies (Consistent with the Paris Agreement)," Discussion paper ES 17-6, Harvard Project on Climate Agreements, October 2017.
4. World Bank, Ecofys, and Vivid Economics, *State and Trends of Carbon Pricing 2016* (World Bank, Washington, DC, 2016).
5. UNFCCC, "Paris Agreement, Article 6" (UN, Paris, France, 2015), pp. 7–8.
6. G. E. Metcalf, D. Weisbach, *Rev. Environ. Econ. Policy* **6**, 110 (2012).
7. W. Obergassel, F. Asche, "Shaping the Paris Mechanisms Part III: An Update on Submissions on Article 6 of the Paris Agreement" (Wuppertal Institute for Climate, Environment, and Energy, JIKO Policy Paper 05/2017, 2017).
8. J. Jaffe *et al.*, *Ecol. Law Q.* **36**, 789 (2009).
9. M. Ranson, R. N. Stavins, *Clim. Policy* **16**, 284 (2016).
10. D. M. Bodansky *et al. Clim. Policy* **16**, 956 (2016).
11. M. A. Mehling, "Legal Frameworks for Linking National Emissions Trading Systems," in *The Oxford Handbook of International Climate Change Law*, C. P. Carlarne, K. R. Gray, R. G. Tarasofsky, Eds. (Oxford Univ. Press, 2016), pp. 261–288.

ACKNOWLEDGMENTS

The authors acknowledge comments from S. Biniaz, D. Bodansky, C. Haug, C. Hood, and A. Marcu and financial support from the Enel Foundation.

10.1126/science.aar5988

HISTORY OF SCIENCE

Looking home from the heavens

An ode to early spaceflight celebrates the transformative effect of viewing Earth from above

The iconic "Blue Marble" photograph places humanity's accomplishments and foibles in context.

PHOTO: NASA

By **Matthew Shindell**

On 17 December 1903, Orville and Wilbur Wright performed the first successful controlled flight of a powered aircraft. Less than 60 years later, on 12 April 1961, the Soviet Union launched Yuri Gagarin aboard the Vostok 1 spacecraft into orbit around Earth, making him the first human in space. The half-century between these two landmarks in the history of flight saw two world wars, a tense cold war between two nuclear superpowers, and the accelerated advancement of the emergent aerospace industry.

It was no secret that war, conflict, and competition were the driving forces behind the state's interest in rockets, even if the scientists and engineers claimed loftier goals. So it is quite remarkable that from these pursuits—from countless military and civilian contracts made in the interest of national defense—came some of the most enduring images of a peaceful and fragile Earth.

The most iconic images—the "Earthrise" photograph taken by astronaut Bill Anders on Christmas Eve 1968 and the "Blue Marble" photograph of the entire Earth taken by Harrison "Jack" Schmitt on 7 December 1972—were products of NASA's Apollo program. By some accounts, these images helped to launch a nascent environmental movement and a new political consciousness within the American counterculture. But according to the men who captured them, they hardly do justice to the experience of seeing Earth from a distance, which some have described as a truly spiritual experience.

While Christopher Potter's title—*The Earth Gazers*—and much of the book's thrust focus on the transformative effect of seeing Earth from space, the bulk of it is about the saga of flight, from early aircraft to the Apollo mission. Potter begins his narrative not with the Wright Flyer but with Charles Lindbergh's historic 1927 flight from New York to Paris, presumably because the famous aviator is a more substantial historical figure and there is a wealth of available published material about his life (22 of the 62 sources listed in the book's bibliography are either about Lindbergh or written by a Lindbergh).

Lindbergh, in turn, is used to introduce a number of other highlights of early human flight, providing a narrative thread that runs through the book. We learn, for example, that he encountered and admired the American rocketeer Robert Goddard; that he spent time in Germany during the rise of Hitler and the Nazi party, whose ambitions would propel the development of Wernher von Braun's V-2 rocket; and that he relished the achievements of the Apollo mission, believing the conquest of space to be akin to the pioneering days of flight.

The Earth Gazers
On Seeing Ourselves
Christopher Potter
Pegasus Books, 2018.
470 pp.

The book is split into three parts. In the first, the reader is presented with biographical sketches of many of the already well-known figures in the history of aviation, rocketry, and space exploration, including Goddard, von Braun, and to a lesser extent their Russian counterparts, Konstantin Tsiolkovsky and Sergei Korolev.

As the story turns to the first crewed flights into space, the focus turns to individual missions undertaken by the U.S. space program. Part one ends with a brief history of Project Mercury. Part two extends this story from Gemini to Apollo. Potter profiles each mission and provides anecdotes that bring out the personalities of the astronauts.

Although Potter did not conduct any original research, these first two parts of the book are based on a close reading of the published literature. A reader well versed in the history of air and spaceflight will not learn anything new here, but as a nonacademic introduction to the topic, this book will no doubt be appreciated. Potter's writing style is clear and engaging, and his approach to the subject is thorough, if not scholarly.

As much as this book's narrative approach to the history of human flight deserves praise, it is not clear what the overall message is meant to be. In the final third of the book, Potter emphasizes the transformative effect of seeing the whole Earth from outside, something only a handful of humans have had the privilege to do. He also laments the failure of NASA to convey this experience effectively to the public, suggesting that it was hamstrung by the insistence of some activist organizations that NASA not tread upon religious ground in its depiction of spaceflight—an issue brought to a head in a 1971 lawsuit waged by Madalyn Murray O'Hair, founder of American Atheists, that was spurred by the Christmas Eve reading of the Book of Genesis by the crew of Apollo 8.

Potter's clear narrative style gives way to sometimes opaque poetic musing. He alludes to a "metaphysics of space travel" that so far has been lost on humans. He seems to want humankind to find a new understanding of its place in the universe—one that acknowledges our connectedness with the technologies that now surround us and their potential ability to lift us—and sees the technological triumph of Apollo to literally transcend our earthly bonds as a fitting allegory. ■

The reviewer is space history curator, National Air and Space Museum, Washington, DC 20014, USA. Email: shindellm@si.edu

10.1126/science.aar7911

EVOLUTIONARY BIOLOGY

Evolution's amazing arms race

A friendly tour of unusual animal adaptations misses many highlights

By **Christopher Kemp**

Strange Survivors
How Organisms Attack and Defend in the Game of Life
Oné R. Pagán
BenBella Books, 2018. 240 pp.

The Iberian ribbed newt (*Pleurodeles waltl*) is an evolutionary biologist's dream. Unremarkable in appearance, with rough slate-colored skin, it's found across the Iberian Peninsula and southward in Morocco. Nondescript perhaps, but the newt is superlative. When threatened, its ribs puncture the sides of its body like a row of defensive spears. At the same time, it releases venom that coats its newly exposed ribs.

Elsewhere, evolution has resulted in countless other species with survival strategies as strange as, or stranger than, the Iberian ribbed newt's. West Chester University biologist Oné R. Pagán describes many of them in *Strange Survivors*.

The slow loris produces a potent toxin in glands on its arms.

An expert in flatworms, Pagán introduces us to a variety of creatures well practiced in attacking and defending themselves, from cone snails and jellyfish to poison toads and spitting spiders. There is the slow loris (*Nycticebus* genus), a nocturnal primate from Southeast Asia that produces a two-stage toxin in glands on its arms. When it licks the already-poisonous toxin from its glands, the loris's saliva activates it and makes it more potent still. Like many species, it is unaffected by its own toxin. We meet venomous shrews; mantis shrimp armed with percussively powerful limbs; electric eels, capable of generating 800 volts of electricity; and toxic birds, such as the pitohui of New Guinea, which not only are inedible but also can cause sneezing, stinging, nausea, and numbness when handled.

The first chapter is a primer on evolution; the second provides an outline of important biological concepts such as DNA, metabolism, replication, life, and death. Frustratingly, though, examples of the strange highly evolved species that fill the book's later chapters are few and far between. This is a mistake. It's too much to ask a reader to digest nearly 50 pages of background information with only a glimpse at the organisms alluded to in the book's title.

Scientific writing—even when it is intended for a lay audience—requires precision. Instead of specifics, however, Pagán often resorts to generalities, describing a cell membrane as "rather versatile," a snail's venom-injecting mechanism as "rather sophisticated," and a fish's startle response as "rather useful." The text is punctuated throughout with clichés, another sign of imprecision, including: "to be fair," "come to think of it," "truth be told," and "by the way." The book's many footnotes, several of which send the reader on a detour to the bottom of the page only to find an unnecessary aside about *Star Trek*, are a bit tiring.

Strange Survivors also falls short on a more fundamental level. As far as unusual survival strategies are concerned, the Tree of Life is heavy with fruit. No one would expect a single book to be exhaustive—Pagán himself admits that the book contains but a "small fraction" of unusual survival strategies employed by animals—but he fails to mention so many candidates that eventually one becomes aware of their omission.

There are almost 70 species of flying fish belonging to the *Exocoetidae* family, that, when threatened, leap from the water and fly using long winglike fins, for instance. They can remain airborne for more than 40 seconds and fly for hundreds of feet. But Pagán doesn't mention them.

He doesn't mention the bombardier beetle either, which defends itself with an almost-boiling cloud of noxious gas that explodes from its abdomen. Or the boxer crabs of the *Lybia* genus, which hold a sea anemone in each claw, like colorful pom-poms, to help them catch prey. Or *Octopoteuthis deletron*, a squid that willfully detaches one of its arms to confuse predators.

He doesn't discuss the defensive strategies of the skunk or the porcupine. He doesn't mention mimicry at all, missing an opportunity to highlight the elephant hawk moth (*Deilephila elpenor*) caterpillar, which resembles a snake, and the larvae of the alder moth (*Acronicta alni*), which look like bird droppings.

He doesn't mention camouflage either, neglecting the chameleons and cephalopods that change color so that they match their surroundings; the stick insects that evolved, over millions of years, to become indistinguishable from the tree branches on which they stand; and the stonefish, which resembles, well, a stone.

A better book would have mentioned at least some of these incredible strategies, even briefly, in a bid to illustrate the richness of biodiversity that surrounds us. Regardless, it is clear that Pagán cares deeply for his subject. Even in its incompleteness, his book sent me on a careful search for the creatures he omitted.

As rabbit holes go, it was a pleasant one, riddled with strange and sinuous half-hidden tunnels. Searches like that are filled with meaning. That alone may be enough reason to read *Strange Survivors*. ■

The reviewer is the author of The Lost Species: Great Expeditions in the Collections of Natural History Museums *(University of Chicago Press, Chicago, 2017). Email: cjkemp@gmail.com*

10.1126/science.aar6211

A local resident battles wildfire in Portugal.

PHOTO: GLOBAL MEDIA GROUP/SIPA USA/NEWSCOM

Edited by **Jennifer Sills**

Agricultural policy can reduce wildfires

Last year, once again, forest fires took their toll in southern Europe. In Portugal alone, at least 500,000 ha were burned, 100 people were killed, and 500 houses were lost (*1*, *2*). As in most Mediterranean countries, wildfires raged mainly through abandoned farmland that has turned into forests and shrublands.

Agriculture is an important driver of European wildfires. It is a major source of fire ignitions (*3*, *4*). Additionally, farmland abandonment and policies promoting forestry increase fire hazard, as they lead to vegetation growth and fuel build-up in the landscape (*5*). However, agriculture is also part of the solution. Agricultural areas, such as crops, orchards, and grasslands, are much less fire-prone, particularly if they include irrigated crops (*5*, *6*). The European Union's Common Agricultural Policy (CAP) is a powerful financial instrument that can contribute to sustainable environmental management and climate change adaptation. The vision for CAP, recently proposed by the European Commission (*7*), addresses natural hazards from climate change, including fire, but focuses on farmers and their crops.

The CAP should assume a larger role in reducing fire hazard by addressing four priorities. First, CAP should foster the maintenance or reintroduction of extensive livestock grazing in areas prone to abandonment. Second, CAP should promote agricultural use in the wildland-urban interface, mainly around villages in remote areas where the historical surrounding agricultural area has been lost, resulting in vegetation succession and an increased risk of economic damage and loss of human lives, as fires enter villages. These agricultural belts can passively protect urban areas and valuable infrastructures, in addition to facilitating both firefighting operations and the suppression of fire ignitions. Third, CAP should decrease fire ignitions by regulating the burning of crop residues, the use of fire by shepherds in mountain ranges, and the use of agricultural machinery during the dry season. Fourth, CAP should promote adequate forest management in high–fire risk areas, including protecting and restoring open woodland vegetation (such as wood pastures), giving preference to agro-forestry over dense tree plantations, restoring the use of understory biomass as bio-energy to avoid accumulation of flammable material, and selecting native, less fire-prone, tree species in forestry [such as native oak species instead of pine or eucalyptus (*5*, *8*)].

Megafires are mostly driven by weather conditions (*9*), and with climate change we should anticipate an increase in their frequency and impact, especially in southern Europe (*10*). The current strong investment in fire suppression, in a context of farmland abandonment, results in increased fuel loads and potential for larger future fires (*11*). The European agricultural policy should instead balance fire suppression with nature-based solutions. Multifunctional, fire-resilient, mosaic landscapes can maintain both natural and cultural assets and serve to reduce fire intensity and damage when burned.

Francisco Moreira[1,2]* and Guy Pe'er[3,4,5]
[1]CIBIO/InBIO, University of Porto, 4485-601, Vairão, Portugal. [2]CEABN/InBIO, School of Agronomy, University of Lisbon, Tapada da Ajuda, 1349-017, Lisboa, Portugal. [3]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 04103, Leipzig, Germany. [4]University of Leipzig, 04109, Leipzig, Germany. [5]UFZ–Helmholtz Centre for Environmental Research, 04318, Leipzig, Germany.
*Corresponding author. Email: fmoreira@cibio.up.pt

REFERENCES

1. S. Gómez-González *et al.*, *Environ. Sci. Pol.* **81**, 104 (2018).
2. "Portugal fires burn 520,000 hectares, nearly 60 percent of EU total" *Reuters* (2017); www.reuters.com/article/us-portugal-fire-area/portugal-fires-burn-520000-hectares-nearly-60-percent-of-eu-total-idUSKBN1CN2F4.
3. A. Ganteaume *et al.*, *Environ. Management* **51**, 651 (2013).
4. F. X. Catry, F. C. Rego, F. L. Bação, F. Moreira, *Int. J. Wildland Fire* **18**, 921 (2009).
5. F. Moreira, *J. Environ. Management* **92**, 2389 (2011).
6. S. Oliveira, F. Moreira, R. Boca, J. San-Miguel-Ayanz, J. Pereira, *Int. J. Wildland Fire* **23**, 620 (2013).
7. European Commission, "The future of food and farming: Communication from the Commission to the European Parliament" (European Economic and Social Committee and the Committee of the Regions, 2017).
8. F. Moreira, P. Vaz, F. Catry, J. Silva, *Int. J. Wildland Fire* **18**, 563 (2009).
9. P. M. Fernandes, A. M. G. Barros, A. Pinto, J. A. Santos, *J. Geophys. Res. Biogeosci.* **121**, 2141 (2016).
10. W. M. Jolly *et al.*, *Nat. Commun.* **6**, 7537 (2015).
11. R. D. Collins, R. de Neufville, J. Claro, T. Oliveira, A. P. Pacheco, *J. Environ. Management* **130**, 1 (2013).

10.1126/science.aat1359

Rethinking wildfires and forest watersheds

In December 2017, wildfires burned large swaths of southern California, dramatically ending an already destructive wildfire season in the United States. The 2017 wildfire season burned more than 3.9 million hectares in the United States, the third-most area burned in 1 year since 1960 (*1*). The largest of the fires, the Thomas Fire in Ventura County, CA, burned more than 1140 km^2, including thousands of structures, forcing more than 100,000 residents from their homes (*2*).

The devastating impacts of the most recent wildfire season are consistent with the trends of increasing occurrence of large wildfire activity, longer wildfire durations, and longer wildfire seasons that have been evident since the mid-1980s (*3*). Similar trends of increasing wildfire activity have occurred elsewhere in the world, including

Canada, Australia, and regions of South America, Eurasia, and Africa (*4*). Given that Earth's climate continues to warm and that historical land use and fire suppression activities have resulted in dense forests that provide fuel for fires, these accelerating trends are projected to continue into the foreseeable future (*5*).

The costs associated with fighting these large wildfires now account for more than half of the U.S. Forest Service annual budget. Even before the December wildfires, 2017 was the most expensive year on record, with costs for wildland fire suppression exceeding $2 billion (*6*). However, the full economic costs of wildfire should also consider expenditures associated with preparedness, property losses, health care and loss of human life, tourism, and damage to the natural resource base. The true costs of the fires are likely 2 to 30 times as high as the reported suppression costs (*7*).

Counterintuitively, the threats and costs once fires are contained may be more disastrous than the fire itself. The secondary threats of wildfires to water supply are particularly concerning, as almost two-thirds of municipalities in North America receive their drinking water from forested areas (*8*). Key threats include increased potential for erosion, landslides, debris flows, floods, and introduction of contaminants to streams, with potentially catastrophic implications for community infrastructure, drinking water treatment, public health, and aquatic ecosystem health (*9*).

Given the rising threats and costs associated with the current wildfire trend, we must change the way we manage both wildfires and forested watersheds. For example, the use of prescribed fire or fostering of fires that burn more frequently and under less extreme conditions can improve forest resilience and reduce the magnitude and longevity of effects (*10*). These land-use activities, especially in forests near communities, have potential to substantially reduce impacts if they are strategically located (*11*). However, it is not economical or feasible to protect all forests through active forest management. As such, it is critical to continue to develop and use the tools we have to produce maps that identify locations and times (e.g., early warning systems) of high fire risk, which can guide our policy and management efforts. Such efforts should also integrate and focus on areas that are critical for provision of a freshwater supply, to protect water resources for healthy aquatic ecosystems and human populations downstream (*12*).

Kevin D. Bladon
Oregon State University, Corvallis, OR 97331, USA.
Email: bladonk@oregonstate.edu

Canada's iconic wilderness includes Bow Lake and Crowfoot Mountain in Banff National Park.

REFERENCES

1. National Interagency Fire Centre, "Total wildland fires and acres (1960–2016)" (Boise, ID, 2017).
2. CAL FIRE, "Thomas Fire" (Sacramento, CA, 2018).
3. A. L. Westerling, H. G. Hidalgo, D. R. Cayan, T. W. Swetnam, *Science* **313**, 940 (2006).
4. W. M. Jolly *et al.*, *Nat. Commun.* **6**, 7537 (2015).
5. M. D. Flannigan, M. A. Krawchuk, W. J. de Groot, B. M. Wotton, L. M. Gowman, *Int. J. Wildland Fire* **18**, 483 (2009).
6. U.S. Department of Agriculture, "Forest Service wildland fire suppression costs exceed $2 billion" (Washington, DC, 2017).
7. L. Dale, "The true cost of wildfire in the western U.S." (Western Forestry Leadership Coalition, Lakewood, CO, 2009).
8. Committee on Hydrologic Impacts of Forest Management, "Hydrologic effects of a changing forest landscape" (The National Academies Press, Washington, DC, 2008).
9. K. D. Bladon, M. B. Emelko, U. Silins, M. Stone, *Environ. Sci. Technol.* **48**, 8936 (2014).
10. C. S. Stevens-Rumann *et al.*, *Ecol. Lett.* **21**, 243 (2018).
11. D. M. J. S. Bowman *et al.*, *Nat. Ecol. Evol.* **1**, 0058 (2017).
12. F.-N. Robinne *et al.*, *Sci. Total Environ.* **610-611**, 1193 (2018).

10.1126/science.aar8120

Invest long term in Canada's wilderness

Increasing global demand for Canada's resources is eroding the country's iconic wilderness, intact ecosystems, and rich megafaunal diversity (*1*, *2*). To meet its 2020 commitments to the United Nations Convention on Biological Diversity (CBD), Canada must protect 17% of its terrestrial area and 10% of its marine area (*3*); currently, only 10 and 1%, respectively, are protected (*4*). Polls suggest that 87% of Canadians support increased landscape protection (*5*). On 8 January, 116 Canadian politicians called for a historic $1.4 billion in government funding to conserve Canada's exceptional wilderness and biodiversity between 2018 and 2020, with $470 million per year to support efforts after 2020 (*3*). This investment is essential to enact the land and water protection Canadians want. We support this call to action.

However, even if Canada meets its CBD commitment to protect 17% of its terrestrial area, wildlife conservation will fail if Canada neglects the other 83%, which will remain unprotected. In western Canada, 35% of the provincially managed landscape has been affected by industrial activity (*6*). These effects are gradually compromising the persistence of many high-profile species, including the grizzly bear, caribou, elk, wolverine, and mountain goat (*6*). The growing threats to Canada's functional ecosystems are not matched by increasing funds to manage and conserve wildlife and habitats. Funds provided to wildlife management agencies in western Canada pale in comparison to neighboring jurisdictions and are in decline (*7*). We strongly urge provincial governments to honor their promise to address this wide funding deficit (*8*) to ensure the effective management and conservation of Canada's species outside protected areas.

Canadian governments have a responsibility not only to their citizens, who overwhelmingly support conservation, but also to the world as stewards of 24% of the planet's remaining wilderness (*2*). Increased investment in both protected and unprotected areas is vital to safeguard Canada's immense wilderness and wildlife capital.

Clayton T. Lamb,[1]* Marco Festa-Bianchet,[2] Mark S. Boyce[1]
[1]Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2E9, Canada. [2]Université de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada.
*Corresponding author. Email: ctlamb@ualberta.ca

REFERENCES

1. M. Hebblewhite, *Biol. Cons.* **206**, 102 (2017).
2. J. Allan, O. Venter, J. Watson, *Sci. Data* **4**, 1 (2017).
3. W. Amos *et al.*, "Support for a historic investment in protecting Canada's land, freshwater, and ocean" (2018); http://wamos.liberal.ca/wp-content/uploads/sites/1508/2018/01/EN_Final-letter-with-signatories-Budget-2018-1-1.pdf.
4. Environment and Climate Change Canada, "Canadian environmental sustainability indicators: Canada's protected areas" (2017); www.ec.gc.ca/indicateurs-indicators/default.asp?lang=en&n=478A1D3D-1.
5. Earnscliffe Strategy Group, "National conservation survey" (2017); http://earnscliffe.ca/wp-content/uploads/2017/11/National-Conservation-Survey.pdf.
6. N. Shackelford, R. J. Standish, W. Ripple, B. M. Starzomski, *Cons. Biol.* 10.1111/cobi.13036 (2017).
7. R. Archibald, D.S. Eastman, R. Ellis, B. Nyberg, *J. Ecosyst. Management* **14**, 1 (2014).
8. "British Columbia to increase investment in wildlife management," *BC Gov News* (2017); https://news.gov.bc.ca/14367.

10.1126/science.aat1104

A local resident battles wildfire in Portugal.

PHOTO: GLOBAL MEDIA GROUP/SIPA USA/NEWSCOM

Edited by **Jennifer Sills**

Agricultural policy can reduce wildfires

Last year, once again, forest fires took their toll in southern Europe. In Portugal alone, at least 500,000 ha were burned, 100 people were killed, and 500 houses were lost (*1*, *2*). As in most Mediterranean countries, wildfires raged mainly through abandoned farmland that has turned into forests and shrublands.

Agriculture is an important driver of European wildfires. It is a major source of fire ignitions (*3*, *4*). Additionally, farmland abandonment and policies promoting forestry increase fire hazard, as they lead to vegetation growth and fuel build-up in the landscape (*5*). However, agriculture is also part of the solution. Agricultural areas, such as crops, orchards, and grasslands, are much less fire-prone, particularly if they include irrigated crops (*5*, *6*). The European Union's Common Agricultural Policy (CAP) is a powerful financial instrument that can contribute to sustainable environmental management and climate change adaptation. The vision for CAP, recently proposed by the European Commission (*7*), addresses natural hazards from climate change, including fire, but focuses on farmers and their crops.

The CAP should assume a larger role in reducing fire hazard by addressing four priorities. First, CAP should foster the maintenance or reintroduction of extensive livestock grazing in areas prone to abandonment. Second, CAP should promote agricultural use in the wildland-urban interface, mainly around villages in remote areas where the historical surrounding agricultural area has been lost, resulting in vegetation succession and an increased risk of economic damage and loss of human lives, as fires enter villages. These agricultural belts can passively protect urban areas and valuable infrastructures, in addition to facilitating both firefighting operations and the suppression of fire ignitions. Third, CAP should decrease fire ignitions by regulating the burning of crop residues, the use of fire by shepherds in mountain ranges, and the use of agricultural machinery during the dry season. Fourth, CAP should promote adequate forest management in high–fire risk areas, including protecting and restoring open woodland vegetation (such as wood pastures), giving preference to agro-forestry over dense tree plantations, restoring the use of understory biomass as bio-energy to avoid accumulation of flammable material, and selecting native, less fire-prone, tree species in forestry [such as native oak species instead of pine or eucalyptus (*5*, *8*)].

Megafires are mostly driven by weather conditions (*9*), and with climate change we should anticipate an increase in their frequency and impact, especially in southern Europe (*10*). The current strong investment in fire suppression, in a context of farmland abandonment, results in increased fuel loads and potential for larger future fires (*11*). The European agricultural policy should instead balance fire suppression with nature-based solutions. Multifunctional, fire-resilient, mosaic landscapes can maintain both natural and cultural assets and serve to reduce fire intensity and damage when burned.

Francisco Moreira[1,2]* and Guy Pe'er[3,4,5]
[1]CIBIO/InBIO, University of Porto, 4485-601, Vairão, Portugal. [2]CEABN/InBIO, School of Agronomy, University of Lisbon, Tapada da Ajuda, 1349-017, Lisboa, Portugal. [3]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 04103, Leipzig, Germany. [4]University of Leipzig, 04109, Leipzig, Germany. [5]UFZ–Helmholtz Centre for Environmental Research, 04318, Leipzig, Germany.
*Corresponding author. Email: fmoreira@cibio.up.pt

REFERENCES

1. S. Gómez-González *et al.*, *Environ. Sci. Pol.* **81**, 104 (2018).
2. "Portugal fires burn 520,000 hectares, nearly 60 percent of EU total" *Reuters* (2017); www.reuters.com/article/us-portugal-fire-area/portugal-fires-burn-520000-hectares-nearly-60-percent-of-eu-total-idUSKBN1CN2F4.
3. A. Ganteaume *et al.*, *Environ. Management* **51**, 651 (2013).
4. F. X. Catry, F. C. Rego, F. L. Bação, F. Moreira, *Int. J. Wildland Fire* **18**, 921 (2009).
5. F. Moreira, *J. Environ. Management* **92**, 2389 (2011).
6. S. Oliveira, F. Moreira, R. Boca, J. San-Miguel-Ayanz, J. Pereira, *Int. J. Wildland Fire* **23**, 620 (2013).
7. European Commission, "The future of food and farming: Communication from the Commission to the European Parliament" (European Economic and Social Committee and the Committee of the Regions, 2017).
8. F. Moreira, P. Vaz, F. Catry, J. Silva, *Int. J. Wildland Fire* **18**, 563 (2009).
9. P. M. Fernandes, A. M. G. Barros, A. Pinto, J. A. Santos, *J. Geophys. Res. Biogeosci.* **121**, 2141 (2016).
10. W. M. Jolly *et al.*, *Nat. Commun.* **6**, 7537 (2015).
11. R. D. Collins, R. de Neufville, J. Claro, T. Oliveira, A. P. Pacheco, *J. Environ. Management* **130**, 1 (2013).

10.1126/science.aat1359

Rethinking wildfires and forest watersheds

In December 2017, wildfires burned large swaths of southern California, dramatically ending an already destructive wildfire season in the United States. The 2017 wildfire season burned more than 3.9 million hectares in the United States, the third-most area burned in 1 year since 1960 (*1*). The largest of the fires, the Thomas Fire in Ventura County, CA, burned more than 1140 km^2, including thousands of structures, forcing more than 100,000 residents from their homes (*2*).

The devastating impacts of the most recent wildfire season are consistent with the trends of increasing occurrence of large wildfire activity, longer wildfire durations, and longer wildfire seasons that have been evident since the mid-1980s (*3*). Similar trends of increasing wildfire activity have occurred elsewhere in the world, including

Canada, Australia, and regions of South America, Eurasia, and Africa (*4*). Given that Earth's climate continues to warm and that historical land use and fire suppression activities have resulted in dense forests that provide fuel for fires, these accelerating trends are projected to continue into the foreseeable future (*5*).

The costs associated with fighting these large wildfires now account for more than half of the U.S. Forest Service annual budget. Even before the December wildfires, 2017 was the most expensive year on record, with costs for wildland fire suppression exceeding $2 billion (*6*). However, the full economic costs of wildfire should also consider expenditures associated with preparedness, property losses, health care and loss of human life, tourism, and damage to the natural resource base. The true costs of the fires are likely 2 to 30 times as high as the reported suppression costs (*7*).

Counterintuitively, the threats and costs once fires are contained may be more disastrous than the fire itself. The secondary threats of wildfires to water supply are particularly concerning, as almost two-thirds of municipalities in North America receive their drinking water from forested areas (*8*). Key threats include increased potential for erosion, landslides, debris flows, floods, and introduction of contaminants to streams, with potentially catastrophic implications for community infrastructure, drinking water treatment, public health, and aquatic ecosystem health (*9*).

Given the rising threats and costs associated with the current wildfire trend, we must change the way we manage both wildfires and forested watersheds. For example, the use of prescribed fire or fostering of fires that burn more frequently and under less extreme conditions can improve forest resilience and reduce the magnitude and longevity of effects (*10*). These land-use activities, especially in forests near communities, have potential to substantially reduce impacts if they are strategically located (*11*). However, it is not economical or feasible to protect all forests through active forest management. As such, it is critical to continue to develop and use the tools we have to produce maps that identify locations and times (e.g., early warning systems) of high fire risk, which can guide our policy and management efforts. Such efforts should also integrate and focus on areas that are critical for provision of a freshwater supply, to protect water resources for healthy aquatic ecosystems and human populations downstream (*12*).

Kevin D. Bladon
Oregon State University, Corvallis, OR 97331, USA.
Email: bladonk@oregonstate.edu

Canada's iconic wilderness includes Bow Lake and Crowfoot Mountain in Banff National Park.

REFERENCES

1. National Interagency Fire Centre, "Total wildland fires and acres (1960–2016)" (Boise, ID, 2017).
2. CAL FIRE, "Thomas Fire" (Sacramento, CA, 2018).
3. A. L. Westerling, H. G. Hidalgo, D. R. Cayan, T. W. Swetnam, *Science* **313**, 940 (2006).
4. W. M. Jolly *et al.*, *Nat. Commun.* **6**, 7537 (2015).
5. M. D. Flannigan, M. A. Krawchuk, W. J. de Groot, B. M. Wotton, L. M. Gowman, *Int. J. Wildland Fire* **18**, 483 (2009).
6. U.S. Department of Agriculture, "Forest Service wildland fire suppression costs exceed $2 billion" (Washington, DC, 2017).
7. L. Dale, "The true cost of wildfire in the western U.S." (Western Forestry Leadership Coalition, Lakewood, CO, 2009).
8. Committee on Hydrologic Impacts of Forest Management, "Hydrologic effects of a changing forest landscape" (The National Academies Press, Washington, DC, 2008).
9. K. D. Bladon, M. B. Emelko, U. Silins, M. Stone, *Environ. Sci. Technol.* **48**, 8936 (2014).
10. C. S. Stevens-Rumann *et al.*, *Ecol. Lett.* **21**, 243 (2018).
11. D. M. J. S. Bowman *et al.*, *Nat. Ecol. Evol.* **1**, 0058 (2017).
12. F.-N. Robinne *et al.*, *Sci. Total Environ.* **610-611**, 1193 (2018).

10.1126/science.aar8120

Invest long term in Canada's wilderness

Increasing global demand for Canada's resources is eroding the country's iconic wilderness, intact ecosystems, and rich megafaunal diversity (*1*, *2*). To meet its 2020 commitments to the United Nations Convention on Biological Diversity (CBD), Canada must protect 17% of its terrestrial area and 10% of its marine area (*3*); currently, only 10 and 1%, respectively, are protected (*4*). Polls suggest that 87% of Canadians support increased landscape protection (*5*). On 8 January, 116 Canadian politicians called for a historic $1.4 billion in government funding to conserve Canada's exceptional wilderness and biodiversity between 2018 and 2020, with $470 million per year to support efforts after 2020 (*3*). This investment is essential to enact the land and water protection Canadians want. We support this call to action.

However, even if Canada meets its CBD commitment to protect 17% of its terrestrial area, wildlife conservation will fail if Canada neglects the other 83%, which will remain unprotected. In western Canada, 35% of the provincially managed landscape has been affected by industrial activity (*6*). These effects are gradually compromising the persistence of many high-profile species, including the grizzly bear, caribou, elk, wolverine, and mountain goat (*6*). The growing threats to Canada's functional ecosystems are not matched by increasing funds to manage and conserve wildlife and habitats. Funds provided to wildlife management agencies in western Canada pale in comparison to neighboring jurisdictions and are in decline (*7*). We strongly urge provincial governments to honor their promise to address this wide funding deficit (*8*) to ensure the effective management and conservation of Canada's species outside protected areas.

Canadian governments have a responsibility not only to their citizens, who overwhelmingly support conservation, but also to the world as stewards of 24% of the planet's remaining wilderness (*2*). Increased investment in both protected and unprotected areas is vital to safeguard Canada's immense wilderness and wildlife capital.

Clayton T. Lamb,[1]* Marco Festa-Bianchet,[2] Mark S. Boyce[1]
[1]Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2E9, Canada. [2]Université de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada.
*Corresponding author. Email: ctlamb@ualberta.ca

REFERENCES

1. M. Hebblewhite, *Biol. Cons.* **206**, 102 (2017).
2. J. Allan, O. Venter, J. Watson, *Sci. Data* **4**, 1 (2017).
3. W. Amos *et al.*, "Support for a historic investment in protecting Canada's land, freshwater, and ocean" (2018); http://wamos.liberal.ca/wp-content/uploads/sites/1508/2018/01/EN_Final-letter-with-signatories-Budget-2018-1-1.pdf.
4. Environment and Climate Change Canada, "Canadian environmental sustainability indicators: Canada's protected areas" (2017); www.ec.gc.ca/indicateurs-indicators/default.asp?lang=en&n=478A1D3D-1.
5. Earnscliffe Strategy Group, "National conservation survey" (2017); http://earnscliffe.ca/wp-content/uploads/2017/11/National-Conservation-Survey.pdf.
6. N. Shackelford, R. J. Standish, W. Ripple, B. M. Starzomski, *Cons. Biol.* 10.1111/cobi.13036 (2017).
7. R. Archibald, D.S. Eastman, R. Ellis, B. Nyberg, *J. Ecosyst. Management* **14**, 1 (2014).
8. "British Columbia to increase investment in wildlife management," *BC Gov News* (2017); https://news.gov.bc.ca/14367.

10.1126/science.aat1104

RESEARCH

Successful design of oligomeric transmembrane proteins

Lu et al., *p. 1042*

IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

The Arctic fox molts in the autumn from a dark summer to a white winter coat.

ADAPTATION

Changing coats with the season

Many species of mammals and birds molt from summer brown to winter white coats to facilitate camouflage. Mills *et al*. mapped global patterns of seasonal coat color change across eight species including hares, weasels, and foxes. They found regions where individuals molt to white, brown, and both white and brown winter coats. Greater proportions of the populations molted to white in higher latitudes. Regions where seasonal coat changes are the most variable (molting to both brown and white) may provide resilience against the warming climate. —SNV

Science, this issue p. 1033

TUBERCULOSIS

A trehalose tool for tuberculosis

Tuberculosis is the leading infectious killer worldwide. The prevalence of drug- and multidrug-resistant *Mycobacterium tuberculosis* necessitates more rapid and specific diagnostics. Kamariza *et al*. designed a color-changing dye based on trehalose, a sugar that makes up the outer membrane of *M. tuberculosis*. The dye stained live bacteria within minutes, emitting fluorescence upon incorporation into the hydrophobic mycobacterial membrane. Heat-inactivated bacteria did not fluoresce, and drug-treated bacteria emitted reduced fluorescence. This trehalose-based dye does not require sample washing and emits minimal background fluorescence, potentially making it particularly useful for the rapid detection of metabolically active *M. tuberculosis* in resource-limited environments. —CC

Sci. Transl. Med. **10**, eaam6310 (2018).

GEOLOGY

Mudrocks get a vegetative assist

Mudrocks such as slate and shale are rarely found in stratigraphy older than about 500 million years. McMahon and Davies compiled a large database of mudrock occurrence over the past 3.5 billion years to help assess the origin of this ubiquitous rock type (see the Perspective by Fischer). Mudrocks appeared at the same time as did deep-rooted land plants. The interplay between plants and sedimentary rocks suggests that a change in erosion rate and the chemistry of sediments delivered to the oceans occurred around 500 million years ago. —BG

Science, this issue p. 1022; see also p. 994

MOLECULAR BIOLOGY

Tracking regulatory DNA in action

Cis-regulatory DNA elements such as enhancers and promoters are critical for transcription regulation. Little is known about the relationship between these elements' transcriptional activity and their mobility within the nucleus in living cells. Gu *et al*. developed a strategy to deliver multiple RNAs to guide inactive Cas9 to label these elements. Quantitative measurement of their movement during stem cell differentiation revealed that increased DNA loci mobility correlated with transcriptional activation. —SYM

Science, this issue p. 1050

NEUROPHYSIOLOGY

The proton channel behind sour taste

Although many proteins that form ion channels in cell membranes have been described, none that selectively conduct protons into eukaryotic cells have been identified. Tu *et al*.

CREDITS: (FROM TOP TO BOTTOM) P. LU *ET AL.*; NATURE PICTURE LIBRARY/ALAMY STOCK PHOTO

Published by AAAS

used a genetic screen to pinpoint candidate genes that might encode such a protein from mouse taste receptor cells (see the Perspective by Montell). They identified the known protein otopetrin and showed that it conferred proton conductance when expressed in cultured human cells. Their results indicate that otopetrin may function in sensory recognition of sour (acidic) taste in humans and other organisms. —LBR

Science, this issue p. 1047; see also p. 991

ASTRONOMY

Advancing astronomy, one screen saver at a time

One of the major challenges of modern astronomy is finding important results in the overwhelming volume of acquired data. Clark *et al.* used the computers of tens of thousands of volunteers participating in the Einstein@Home project to search Fermi Large Area Telescope data. Their goal was to discover radio-quiet gamma-ray millisecond pulsars (MSPs). More than 10,000 donated CPU years in this survey led to the discovery of two isolated MSPs, including a rotation-powered MSP that has remained undetected in radio observations. —WSW

Sci. Adv. 10.1126/sciadv.aao7228 (2018).

MEDICAL ROBOTS

Individually optimized exosuit

Like a regular suit, a wearable exosuit can be customized for the individual. Ding *et al.* tested a hip-extension device on eight adult male volunteers. The device reduced the energy required to walk, and the authors' approach tailored the assistance for each wearer. They applied a Bayesian algorithm that could rapidly find optimal settings. Optimization reduced the metabolic cost of walking by 17.4%, improving on existing devices by more than 60%. Optimal settings varied between the participants, suggesting that customizing assistance is better than a one-size-fits-all approach. —RLK

Sci. Robot. **3**, eaar5438 (2018).

ORGANIC CHEMISTRY

Guiding nitrenes away from a migration

Nitrogen conventionally shares its electrons in three bonds with one or more partners. A singly bonded nitrogen, or nitrene, is exceptionally reactive and can insert itself into normally inert C–H bonds. If the nitrene forms next to a carbonyl center, though, it tends to react with the C–C bond on the other side instead. Hong *et al*. used theory to guide the design of an iridium catalyst that inhibits this rearrangement, steering the nitrene toward C–H insertion to form a variety of useful lactam rings. —JSY

Science, this issue p. 1016

EVOLUTION

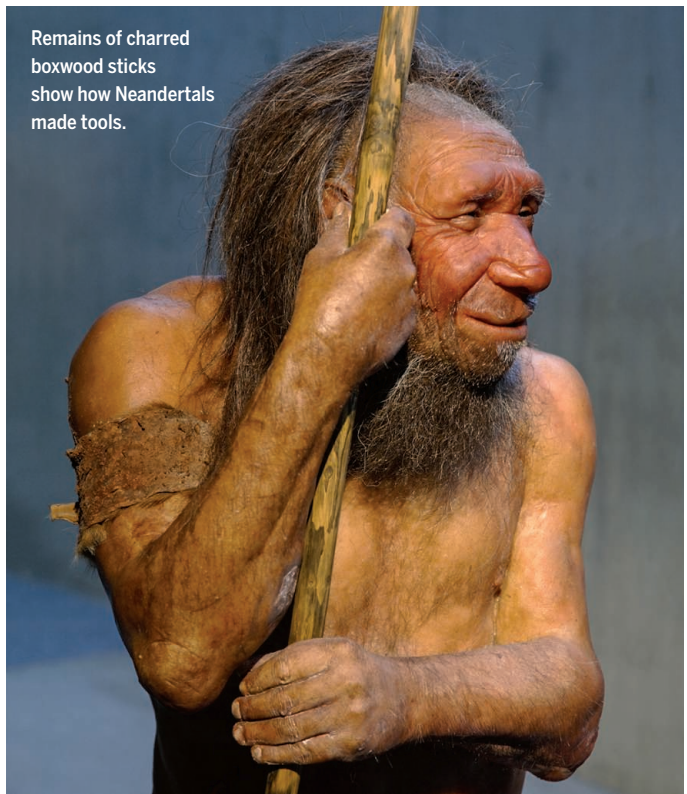
Ratcheting up wild virulence

Partially protective vaccination can sometimes select for increasingly virulent pathogens. Fleming-Davies *et al*. asked what happens in a natural system. In the United States, the house finch population is suffering an increasingly virulent epidemic caused by *Mycoplasma gallisepticum*. The pathogen induces incomplete immunity that clears less virulent pathogens and offers partial protection against strains of greater virulence. In the birds, the partial immune response does away with competition from the less virulent pathogens. The partial immunity of the host also hinders replication of the more virulent pathogens enough to allow some birds to survive. This allows increasingly virulent forms of the pathogen to be transmitted. —CA

Science, this issue p. 1030

IN OTHER JOURNALS

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



Remains of charred boxwood sticks show how Neandertals made tools.

ANTHROPOLOGY

Ablaze in Pleistocene Italy

As engravers know, boxwood is dense and hard. Neandertals knew this too. Nevertheless, wooden artifacts are vulnerable to decay, and such finds are rare and exciting. During excavations for a spa in central Italy, Aranguren *et al.* found remains of elephants, together with remnants of more than 50 burnt wooden sticks dating from around 170,000 years ago. Back then, this area consisted of patches of hotspring wetlands surrounded by grass and box shrub (*Buxus sempervirens*), through which elephant and deer roamed. The Neandertals apparently selected boxwood for its hardness and charred it to make shaping of the tough wood with flint tools a little easier. The tools appear to be of multipurpose digging sticks with rounded handles and pointed tips. Tools of similar dimensions and technology were, until recently, also part of the essential equipment of modern hunter-gatherers. —CA

Proc. Natl. Acad. Sci. U.S.A. 10.1073/pnas.1716068115 (2018).

IMMUNOLOGY

Commensals direct wound healing

Human skin is coated with microorganisms, some of which can cause serious infection when the skin barrier is wounded. But not always. Microorganisms on all barrier surfaces are continually monitored by the host's various immune responses. Linehan *et al.* found that the skin-dwelling organism *Staphylococcus epidermidis* specifically prompts

PHOTO: CARO/RUPERT OBERHAEUSER/NEWSCOM

ALSO IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

SCIENCE COMMUNITY

The whys and wherefores of SciSci

The science of science (SciSci) is based on a transdisciplinary approach that uses large data sets to study the mechanisms underlying the doing of science—from the choice of a research problem to career trajectories and progress within a field. In a Review, Fortunato *et al.* explain that the underlying rationale is that with a deeper understanding of the precursors of impactful science, it will be possible to develop systems and policies that improve each scientist's ability to succeed and enhance the prospects of science as a whole. —BJ

Science, this issue p. 1007

NEUROSCIENCE

How to select and shape neural activity

When we learn a new skill or task, our movements are reinforced and shaped. Learning occurs because the neural activity patterns in the movement control–related brain regions that are rewarded are repeated. But how does this reinforcement work? Athalye *et al*. developed a closed-loop self-stimulation paradigm in which a target motor cortical activity pattern resulted in the optogenetic stimulation of dopaminergic neurons. With training, mice learned to reenter specific neuronal activity patterns, which triggered self-stimulation and shaped their neural activity to be closer to the target pattern. —PRS

Science, this issue p. 1024

MICROBIOLOGY

Maps of defense arsenals in microbial genomes

To survive the attack of foreign invaders such as viruses and plasmids, bacteria and archaea fight back with immune systems that are usually clustered in "defense islands" in their genomes. Doron *et al*. took advantage of this property to map microbial defense systems systematically (see the Perspective by Kim). Candidate immune systems were then experimentally validated for their activities. Like well-known defense arsenals such as restriction-modification and CRISPR systems, these additional immune systems now require mechanistic investigation and could potentially be engineered into useful molecular tools in the future. —SYM

Science, this issue p. 1008; see also p. 993

OPTICS

Exploring photonic topology

Scattering topological effects are being explored in a variety of electronic and optical materials systems owing to their robustness against defects (see the Perspective by Özdemir). Yang *et al.* designed and fabricated an ideal optical analog of a three-dimensional Weyl system. Angular transmission measurements revealed four Weyl points at the same energy, as well as the signature helicoidal arcs associated with such an exotic topological system. Zhou *et al.* theoretically proposed and experimentally demonstrated the formation of a topologically protected bulk Fermi arc. They attributed the formation of the arc to the topological nature of paired exceptional points (points at which gain and loss in the system are matched). Photonic crystals may provide a powerful platform for studying exotic properties of topological electronic systems and may also be used to develop optical devices that exploit topological properties of light-matter interactions. —ISO

Science, this issue p. 1013, p. 1009; see also p. 995

PROTEIN DESIGN

Membrane protein oligomers by design

In recent years, soluble protein design has achieved successes such as artificial enzymes and large protein cages. Membrane proteins present a considerable design challenge, but here too there have been advances, including the design of a zinc-transporting tetramer. Lu *et al*. report the design of stable transmembrane monomers, homodimers, trimers, and tetramers with up to eight membrane-spanning regions in an oligomer. The designed proteins adopted the target oligomerization state and localized to the predicted cellular membranes, and crystal structures of the designed dimer and tetramer reflected the design models. —VV

Science, this issue p. 1042

NEUROIMMUNOLOGY

An off switch for helminth immunity

Group 2 innate lymphoid cells (ILC2s) are involved in responses to helminths, viruses, and allergens. Moriyama *et al*. found that ILC2s interact with the nervous system to modulate helminth immunity. ILC2s from the small intestine expressed the β_2-adrenergic receptor (β_2AR), which normally interacts with the neurotransmitter epinephrine. Inactivating β_2AR resulted in lower helminth burden and more ILC2s, eosinophils, and type 2 cytokine production in mice. Conversely, treatment of helminth-infected mice with a β_2AR agonist enhanced worm burden and reduced proliferation of ILC2s. Thus, β_2AR negatively regulates ILC2-driven protective immunity. —PNK

Science, this issue p. 1056

IMMUNE ENGINEERING

Engineering cytokine-receptor pairs

Interleukin-2 (IL-2) is an important cytokine that helps T cells destroy tumors and virus-infected cells. IL-2 has great therapeutic promise but is limited by toxic side effects and its capacity to both activate and repress immune responses. Sockolosky *et al*. set out to improve IL-2–based immunotherapy by engineering synthetic IL-2–receptor pairs (i.e., IL-2 and its receptor, IL-2R) (see the Perspective by Mackall). Engineered complexes transmitted IL-2 signals but only interacted with each other and not with endogenous IL-2/IL-2R. Treatment of mice with IL-2 improved the ability of engineered T cells to reject tumors with no obvious side effects. This type of approach may provide a way to mitigate toxicities associated with some cytokine-based immunotherapies. —PNK

Science, this issue p. 1037; see also p. 990

ECOLOGY

Fostering the resilience of ecosystems

What makes an ecosystem resilient to environmental change? Answering this question is increasingly important as climate change and other stressors affect ecosystems around the world. But, as Willis *et al*. explain in a Perspective, identifying resilient ecosystems and determining the factors behind their resilience can be difficult. Focusing on tropical ecosystems, the authors outline the many biotic and abiotic factors contributing to resilience and highlight recent studies that provide pointers for future research and conservation aimed at fostering resilience. In another Perspective, Darling and Côté discuss whether coral reef ecosystems might be

protected through conservation approaches and technological interventions focused on increasing their resilience. —JFU

Science, this issue p. 988, p. 986

ALLERGY

Dectin-1 limits allergic responses

Aberrant activation of pattern recognition receptors (PRRs) drives inflammation in autoimmune and allergic diseases. Gour *et al*. identified invertebrate tropomyosin from house dust mites and shrimp as a ligand for dectin-1. Dectin-1 is a PRR that recognizes fungal β-glucans in antifungal immune responses. Engagement of dectin-1 by invertebrate tropomyosins limited type 2 inflammation, and dectin-1–deficient mice were more prone to allergic airway inflammation. Furthermore, expression of dectin-1 was repressed in allergic individuals. Thus, dectin-1 is important in limiting allergic responses. —AB

Sci. Immunol. **3**, eaam9841 (2018).

INFLAMMATION

Tracking inflammation in the colon

The therapeutic options for ulcerative colitis, a form of inflammatory bowel disease, are limited. Lyons *et al*. tracked gene expression, protein levels, and protein phosphorylation in individual animals in a mouse model of colitis. Computational analysis of the data sets identified discrepancies between transcriptomic and proteomic measurements and predicted that the kinase Pak1 mediated colonic inflammation. Treatment of mice with a pharmacological inhibitor of Pak1 ameliorated disease, highlighting the importance of proteomic measurement to the understanding of disease pathogenesis. —JFF

Sci. Signal. **11**, eaan3580 (2018).

used a genetic screen to pinpoint candidate genes that might encode such a protein from mouse taste receptor cells (see the Perspective by Montell). They identified the known protein otopetrin and showed that it conferred proton conductance when expressed in cultured human cells. Their results indicate that otopetrin may function in sensory recognition of sour (acidic) taste in humans and other organisms. —LBR

Science, this issue p. 1047; see also p. 991

ASTRONOMY

Advancing astronomy, one screen saver at a time

One of the major challenges of modern astronomy is finding important results in the overwhelming volume of acquired data. Clark *et al.* used the computers of tens of thousands of volunteers participating in the Einstein@Home project to search Fermi Large Area Telescope data. Their goal was to discover radio-quiet gamma-ray millisecond pulsars (MSPs). More than 10,000 donated CPU years in this survey led to the discovery of two isolated MSPs, including a rotation-powered MSP that has remained undetected in radio observations. —WSW

Sci. Adv. 10.1126/sciadv.aao7228 (2018).

MEDICAL ROBOTS

Individually optimized exosuit

Like a regular suit, a wearable exosuit can be customized for the individual. Ding *et al.* tested a hip-extension device on eight adult male volunteers. The device reduced the energy required to walk, and the authors' approach tailored the assistance for each wearer. They applied a Bayesian algorithm that could rapidly find optimal settings. Optimization reduced the metabolic cost of walking by 17.4%, improving on existing devices by more than 60%. Optimal settings varied between the participants, suggesting that customizing assistance is better than a one-size-fits-all approach. —RLK

Sci. Robot. **3**, eaar5438 (2018).

ORGANIC CHEMISTRY

Guiding nitrenes away from a migration

Nitrogen conventionally shares its electrons in three bonds with one or more partners. A singly bonded nitrogen, or nitrene, is exceptionally reactive and can insert itself into normally inert C–H bonds. If the nitrene forms next to a carbonyl center, though, it tends to react with the C–C bond on the other side instead. Hong *et al*. used theory to guide the design of an iridium catalyst that inhibits this rearrangement, steering the nitrene toward C–H insertion to form a variety of useful lactam rings. —JSY

Science, this issue p. 1016

EVOLUTION

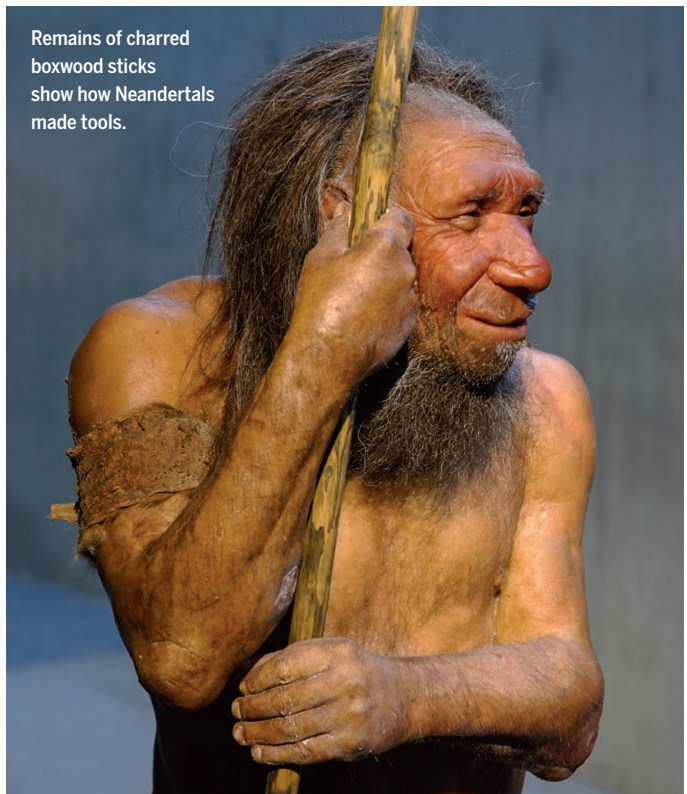
Ratcheting up wild virulence

Partially protective vaccination can sometimes select for increasingly virulent pathogens. Fleming-Davies *et al*. asked what happens in a natural system. In the United States, the house finch population is suffering an increasingly virulent epidemic caused by *Mycoplasma gallisepticum*. The pathogen induces incomplete immunity that clears less virulent pathogens and offers partial protection against strains of greater virulence. In the birds, the partial immune response does away with competition from the less virulent pathogens. The partial immunity of the host also hinders replication of the more virulent pathogens enough to allow some birds to survive. This allows increasingly virulent forms of the pathogen to be transmitted. —CA

Science, this issue p. 1030

IN OTHER JOURNALS

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



Remains of charred boxwood sticks show how Neandertals made tools.

ANTHROPOLOGY

Ablaze in Pleistocene Italy

As engravers know, boxwood is dense and hard. Neandertals knew this too. Nevertheless, wooden artifacts are vulnerable to decay, and such finds are rare and exciting. During excavations for a spa in central Italy, Aranguren *et al.* found remains of elephants, together with remnants of more than 50 burnt wooden sticks dating from around 170,000 years ago. Back then, this area consisted of patches of hotspring wetlands surrounded by grass and box shrub (*Buxus sempervirens*), through which elephant and deer roamed. The Neandertals apparently selected boxwood for its hardness and charred it to make shaping of the tough wood with flint tools a little easier. The tools appear to be of multipurpose digging sticks with rounded handles and pointed tips. Tools of similar dimensions and technology were, until recently, also part of the essential equipment of modern hunter-gatherers. —CA

Proc. Natl. Acad. Sci. U.S.A. 10.1073/pnas.1716068115 (2018).

IMMUNOLOGY

Commensals direct wound healing

Human skin is coated with microorganisms, some of which can cause serious infection when the skin barrier is wounded. But not always. Microorganisms on all barrier surfaces are continually monitored by the host's various immune responses. Linehan *et al.* found that the skin-dwelling organism *Staphylococcus epidermidis* specifically prompts

PHOTO: CARO/RUPERT OBERHAEUSER/NEWSCOM

ASTRONOMY

Why don't astronomers publish observations?

Observing time on top astronomical telescopes is heavily oversubscribed and allocated through a competitive proposal process, but 30 to 50% of observations never produce a peer-reviewed paper. Patat *et al*. contacted principal investigators on proposals executed by the European Southern Observatory (ESO) between 2006 and 2013 that did not result in publication. They found several recurrent issues, such as poor observing weather, lack of investigator resources for analysis, or personnel having left research. By comparing publication rates over time, they showed that 50% of publications occur within 3.5 years of the observations, but it takes 10 years for that fraction to reach 95%. Astronomers are struggling to keep up with the flow of data. —KTS

Messenger **170**, 51 (2017).

The La Silla Observatory in Chile, an ESO facility

an ancient arm of the immune system (the major histocompatibility complex class 1b molecule H2-M3) to respond in a way that avoids inflammation—distinct from responses to pathogens. H2-M3 processes and presents *S. epidermidis*–derived *N*-formyl methionine peptides to $CD8^+$ T cells. These cells express immunoregulatory and tissue-repair gene signatures and accelerate skin wound healing. Hence, hosts and microbiota can interact in highly beneficial ways that may hold promise for therapeutic interventions. —STS

Cell 10.1016/j.cell.2017.12.033 (2018).

IMMUNOGENETICS

Context for immune responses

Immune cells function by recognizing pathogens and initiating a complex cellular response to mount a defense. People can show a wide range of genetically driven variation in responses to infection. Alasoo *et al*. asked how the cellular environment drives change in noncoding regions associated with transcription in immune cells. Chromatin accessibility and gene expression change in pluripotent stem cell lines exposed to signals simulating bacterial infections. Genetic variants among these cell lines affected the timing of gene expression, depending on to what they were exposed. Interestingly, disease-risk variants associated with immune dysfunction, such as rheumatoid arthritis and inflammatory bowel disease, emerged from the analyses. —LMZ

Nat. Genet. 10.1038/s41588-018-0046-7 (2018).

NEURODEVELOPMENT

Regulated tether controls asymmetric cell fate

In the developing fruitfly, the transcription factor Prospero regulates cellular quiescence and fate. Hannaford *et al.* show how an adaptor protein called Miranda manages Prospero activities through sequestration. During interphase in neuroblasts, Miranda binds directly to lipids of the plasma membrane, keeping Prospero away from the cellular nucleus, where it could incite quiescence. During metaphase, thanks to a key phosphorylation, Miranda instead binds to actin bundles at the basal pole of the cell. Prospero, thus tethered, is delivered to one but not the other of the daughter cells, leading to asymmetric allocation of cellular fate. —PJH

eLife **7**, e29939 (2018).

BIOSYNTHESIS

A fitting way to finish

The cofactor heme is used by enzymes to perform challenging oxidations, including reactions on the cofactor itself. Streit *et al.* used spectroscopy to identify a tyrosyl radical formed during the final step of heme biosynthesis. Analysis of the reaction kinetics suggests that the radical is generated by a high-valent iron intermediate created by reaction of hydrogen peroxide with the substrate, which also serves as a cofactor in the reaction. The authors propose that the tyrosyl radical abstracts a hydrogen atom from the substrate, leading to formation of the product, heme b, by oxidative decarboxylation of the propionic acid side chains. —MAF

J. Biol. Chem. 10.1074/jbc.RA117.000830 (2018).

FLUORINE CHEMISTRY

F caught in a bridge

Fluorine typically forms a single bond and then clings to the rest of its electrons with an unrivalled grip. Recently, though, the reactivity of a rigid polycyclic test compound suggested that a fluorine atom could fleetingly bridge two carbon centers in a $(C–F–C)^+$ fluoronium motif. Pitts *et al.* have now spotted that intermediate directly, using a combination of ^{19}F, ^{1}H, and ^{13}C nuclear magnetic resonance spectroscopy. They started at –120°C, but it turned out that the bridge motif was stable up to –40°C. —JSY

Angew. Chem. Int. Ed. **57**, 1924 (2018).

REVIEW SUMMARY

SCIENCE COMMUNITY

Science of science

Santo Fortunato,* Carl T. Bergstrom, Katy Börner, James A. Evans, Dirk Helbing, Staša Milojević, Alexander M. Petersen, Filippo Radicchi, Roberta Sinatra, Brian Uzzi, Alessandro Vespignani, Ludo Waltman, Dashun Wang, Albert-László Barabási*

BACKGROUND: The increasing availability of digital data on scholarly inputs and outputs—from research funding, productivity, and collaboration to paper citations and scientist mobility—offers unprecedented opportunities to explore the structure and evolution of science. The science of science (SciSci) offers a quantitative understanding of the interactions among scientific agents across diverse geographic and temporal scales: It provides insights into the conditions underlying creativity and the genesis of scientific discovery, with the ultimate goal of developing tools and policies that have the potential to accelerate science. In the past decade, SciSci has benefited from an influx of natural, computational, and social scientists who together have developed big data–based capabilities for empirical analysis and generative modeling that capture the unfolding of science, its institutions, and its workforce. The value proposition of SciSci is that with a deeper understanding of the factors that drive successful science, we can more effectively address environmental, societal, and technological problems.

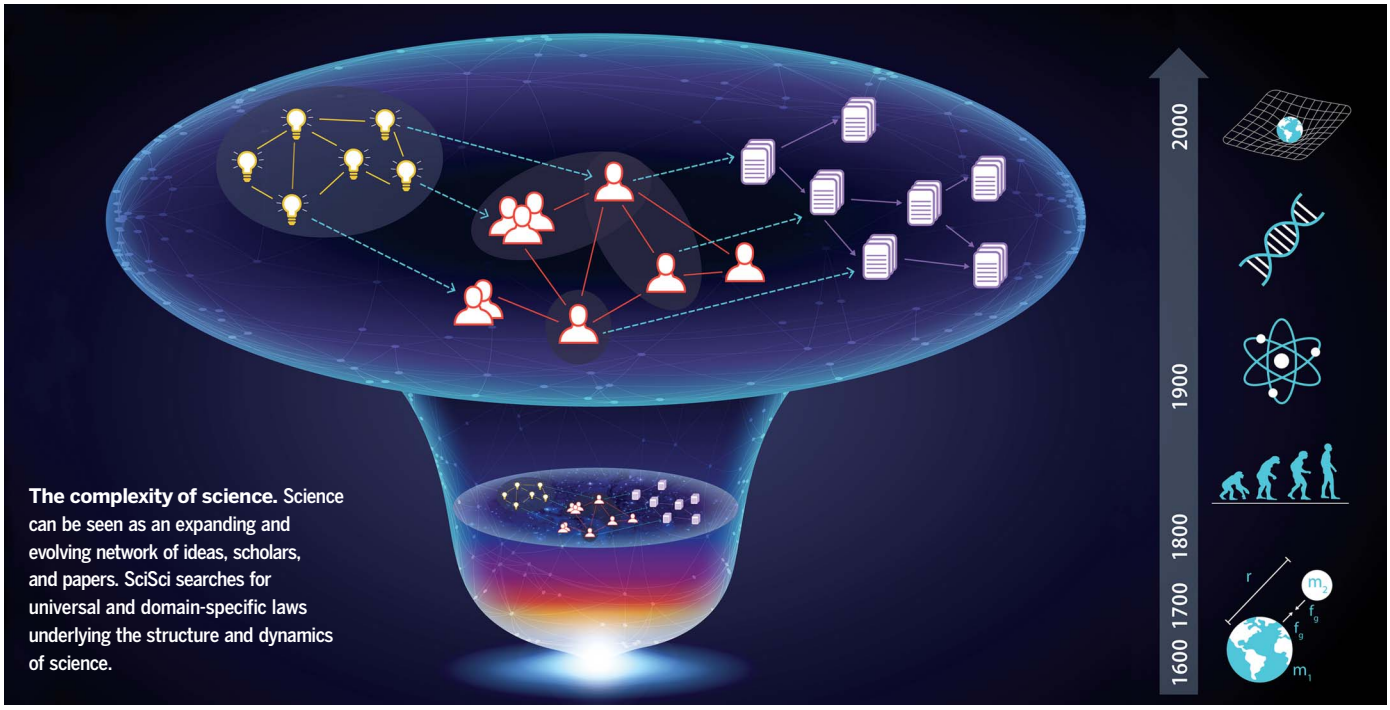
ADVANCES: Science can be described as a complex, self-organizing, and evolving network of scholars, projects, papers, and ideas. This representation has unveiled patterns characterizing the emergence of new scientific fields through the study of collaboration networks and the path of impactful discoveries through the study of citation networks. Microscopic models have traced the dynamics of citation accumulation, allowing us to predict the future impact of individual papers. SciSci has revealed choices and trade-offs that scientists face as they advance both their own careers and the scientific horizon. For example, measurements indicate that scholars are risk-averse, preferring to study topics related to their current expertise, which constrains the potential of future discoveries. Those willing to break this pattern engage in riskier careers but become more likely to make major breakthroughs. Overall, the highest-impact science is grounded in conventional combinations of prior work but features unusual combinations. Last, as the locus of research is shifting into teams, SciSci is increasingly focused on the impact of team research, finding that small teams tend to disrupt science and technology with new ideas drawing on older and less prevalent ones. In contrast, large teams tend to develop recent, popular ideas, obtaining high, but often short-lived, impact.

OUTLOOK: SciSci offers a deep quantitative understanding of the relational structure between scientists, institutions, and ideas because it facilitates the identification of fundamental mechanisms responsible for scientific discovery. These interdisciplinary data-driven efforts complement contributions from related fields such as scientometrics and the economics and sociology of science. Although SciSci seeks long-standing universal laws and mechanisms that apply across various fields of science, a fundamental challenge going forward is accounting for undeniable differences in culture, habits, and preferences between different fields and countries. This variation makes some cross-domain insights difficult to appreciate and associated science policies difficult to implement. The differences among the questions, data, and skills specific to each discipline suggest that further insights can be gained from domain-specific SciSci studies, which model and identify opportunities adapted to the needs of individual research fields. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aao0185

The list of author affiliations is available in the full article online.
*Corresponding author. Email: santo@indiana.edu (S.F.); barabasi@gmail.com (A.-L.B.)
Cite this article as S. Fortunato *et al.*, *Science* **359**, eaao0185 (2018). DOI: 10.1126/science.aao0185



The complexity of science. Science can be seen as an expanding and evolving network of ideas, scholars, and papers. SciSci searches for universal and domain-specific laws underlying the structure and dynamics of science.

ILLUSTRATION: NICOLE SAMAY

REVIEW

SCIENCE COMMUNITY

Science of science

Santo Fortunato,[1,2]* Carl T. Bergstrom,[3] Katy Börner,[2,4] James A. Evans,[5] Dirk Helbing,[6] Staša Milojević,[1] Alexander M. Petersen,[7] Filippo Radicchi,[1] Roberta Sinatra,[8,9,10] Brian Uzzi,[11,12] Alessandro Vespignani,[10,13,14] Ludo Waltman,[15] Dashun Wang,[11,12] Albert-László Barabási[8,10,16]*

Identifying fundamental drivers of science and developing predictive models to capture its evolution are instrumental for the design of policies that can improve the scientific enterprise—for example, through enhanced career paths for scientists, better performance evaluation for organizations hosting research, discovery of novel effective funding vehicles, and even identification of promising regions along the scientific frontier. The science of science uses large-scale data on the production of science to search for universal and domain-specific patterns. Here, we review recent developments in this transdisciplinary field.

The deluge of digital data on scholarly output offers unprecedented opportunities to explore patterns characterizing the structure and evolution of science. The science of science (SciSci) places the practice of science itself under the microscope, leading to a quantitative understanding of the genesis of scientific discovery, creativity, and practice and developing tools and policies aimed at accelerating scientific progress.

The emergence of SciSci has been driven by two key factors. The first is data availability. In addition to the proprietary Web of Science (WoS), the historic first citation index (*1*), multiple data sources are available today (Scopus, PubMed, Google Scholar, Microsoft Academic, the U.S. Patent and Trademark Office, and others). Some of these sources are freely accessible, covering millions of data points pertaining to scientists and their output and capturing research from all over the world and all branches of science. Second, SciSci has benefited from an influx of and collaborations among natural, computational, and social scientists who have developed big data–based capabilities and enabled critical tests of generative models that aim to capture the unfolding of science, its institutions, and its workforce.

One distinctive characteristic of this emerging field is how it breaks down disciplinary boundaries. SciSci integrates findings and theories from multiple disciplines and uses a wide range of data and methods. From scientometrics, it takes the idea of measuring science from large-scale data sources; from the sociology of science, it adopts theoretical concepts and social processes; and from innovation studies, it explores and identifies pathways through which science contributes to invention and economic change. SciSci relies on a broad collection of quantitative methods, from descriptive statistics and data visualization to advanced econometric methods, network science approaches, machine-learning algorithms, mathematical analysis, and computer simulation, including agent-based modeling. The value proposition of SciSci hinges on the hypothesis that with a deeper understanding of the factors behind successful science, we can enhance the prospects of science as a whole to more effectively address societal problems.

Networks of scientists, institutions, and ideas

Contemporary science is a dynamical system of undertakings driven by complex interactions among social structures, knowledge representations, and the natural world. Scientific knowledge is constituted by concepts and relations embodied in research papers, books, patents, software, and other scholarly artifacts, organized into scientific disciplines and broader fields. These social, conceptual, and material elements are connected through formal and informal flows of information, ideas, research practices, tools, and samples. Science can thus be described as a complex, self-organizing, and constantly evolving multiscale network.

Early studies discovered an exponential growth in the volume of scientific literature (*2*), a trend that continues with an average doubling period of 15 years (Fig. 1). Yet, it would be naïve to equate the growth of the scientific literature with the growth of scientific ideas. Changes in the publishing world, both technological and economic, have led to increasing efficiency in the production of publications. Moreover, new publications in science tend to cluster in discrete areas of knowledge (*3*). Large-scale text analysis,

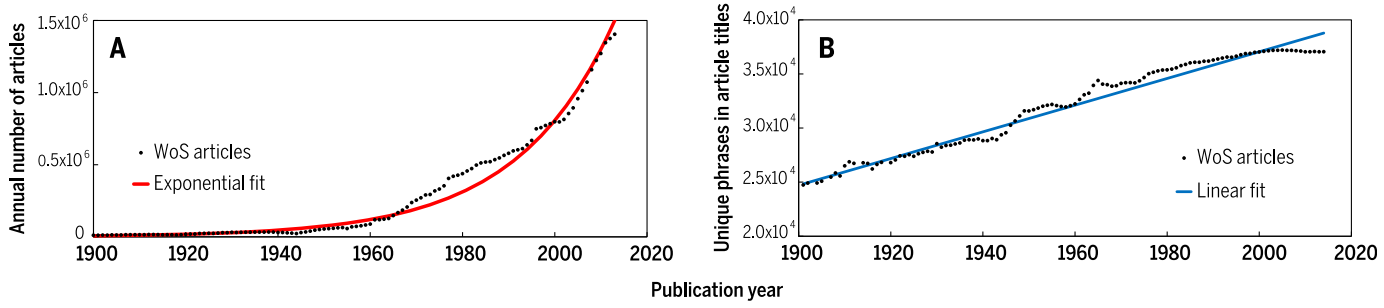


Fig. 1. Growth of science. (**A**) Annual production of scientific articles indexed in the WoS database. (**B**) Growth of ideas covered by articles indexed in the WoS. This was determined by counting unique title phrases (concepts) in a fixed number of articles (*4*).

[1]Center for Complex Networks and Systems Research, School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN 47408, USA. [2]Indiana University Network Science Institute, Indiana University, Bloomington, IN 47408, USA. [3]Department of Biology, University of Washington, Seattle, WA 98195-1800, USA. [4]Cyberinfrastructure for Network Science Center, School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN 47408, USA. [5]Department of Sociology, University of Chicago, Chicago, IL 60637, USA. [6]Computational Social Science, ETH Zurich, Zurich, Switzerland. [7]Ernest and Julio Gallo Management Program, School of Engineering, University of California, Merced, CA 95343, USA. [8]Center for Network Science, Central European University, Budapest 1052, Hungary. [9]Department of Mathematics, Central European University, Budapest 1051, Hungary. [10]Institute for Network Science, Northeastern University, Boston, MA 02115, USA. [11]Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA. [12]Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL 60208, USA. [13]Laboratory for the Modeling of Biological and Sociotechnical Systems, Northeastern University, Boston, MA 02115, USA. [14]ISI Foundation, Turin 10133, Italy. [15]Centre for Science and Technology Studies, Leiden University, Leiden, Netherlands. [16]Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, MA 02115, USA.
*Corresponding author. Email: santo@indiana.edu (S.F.); barabasi@gmail.com (A.-L.B.)

using phrases extracted from titles and abstracts to measure the cognitive extent of the scientific literature, have found that the conceptual territory of science expands linearly with time. In other words, whereas the number of publications grows exponentially, the space of ideas expands only linearly (Fig. 1) (*4*).

Frequently occurring words and phrases in article titles and abstracts propagate via citation networks, punctuated by bursts corresponding to the emergence of new paradigms (*5*). By applying network science methods to citation networks, researchers are able to identify communities as defined by subsets of publications that frequently cite one another (*6*). These communities often correspond to groups of authors holding a common position regarding specific issues (*7*) or working on the same specialized subtopics (*8*). Recent work focusing on biomedical science has illustrated how the growth of the literature reinforces these communities (*9*). As new papers are published, associations (hyperedges) between scientists, chemicals, diseases, and methods ("things," which are the nodes of the network) are added. Most new links fall between things only one or two steps away from each other, implying that when scientists choose new topics, they prefer things directly related to their current expertise or that of their collaborators. This densification suggests that the existing structure of science may constrain what will be studied in the future.

Densification at the boundaries of science is also a signal of transdisciplinary exploration, fusion, and innovation. A life-cycle analysis of eight fields (*10*) shows that successful fields undergo a process of knowledge and social unification that leads to a giant connected component in the collaboration network, corresponding to a sizeable group of regular coauthors. A model in which scientists choose their collaborators through random walks on the coauthorship network successfully reproduces author productivity, the number of authors per discipline, and the interdisciplinarity of papers and authors (*11*).

Problem selection

How do scientists decide which research problems to work on? Sociologists of science have long hypothesized that these choices are shaped by an ongoing tension between productive tradition and risky innovation (*12*, *13*). Scientists who adhere to a research tradition in their domain often appear productive by publishing a steady stream of contributions that advance a focused research agenda. But a focused agenda may limit a researcher's ability to sense and seize opportunities for staking out new ideas that are required to grow the field's knowledge. For example, a case study focusing on biomedical scientists choosing novel chemicals and chemical relationships shows that as fields mature, researchers tend to focus increasingly on established knowledge (*3*). Although an innovative publication tends to result in higher impact than a conservative one, high-risk innovation strategies are rare, because the additional reward does not compensate for the risk of failure to publish at all. Scientific awards and accolades appear to function as primary incentives to resist conservative tendencies and encourage betting on exploration and surprise (*3*). Despite the many factors shaping what scientists work on next, macroscopic patterns that govern changes in research interests along scientific careers are highly reproducible, documenting a high degree of regularity underlying scientific research and individual careers (*14*).

Scientists' choice of research problems affects primarily their individual careers and the careers of those reliant on them. Scientists' collective choices, however, determine the direction of scientific discovery more broadly (Fig. 2). Conservative strategies (*15*) serve individual careers well but are less effective for science as a whole. Such strategies are amplified by the file drawer problem (*16*): Negative results, at odds with established hypotheses, are rarely published, leading to a systemic bias in published research and the canonization of weak and sometimes false facts (*17*). More risky hypotheses may have been tested by generations of scientists, but only those successful enough to result in publications are known to us. One way to alleviate this conservative trap is to urge funding agencies to proactively sponsor risky projects that test truly unexplored hypotheses and take on special interest groups advocating for particular diseases. Measurements show that the allocation of biomedical resources in the United States is more strongly correlated to previous allocations and research than to the actual burden of diseases (*18*), highlighting a systemic misalignment between biomedical needs and resources. This misalignment casts doubts on the degree to which funding agencies, often run by scientists embedded in established paradigms, are likely to influence the evolution of science without introducing additional oversight, incentives, and feedback.

Novelty

Analyses of publications and patents consistently reveal that rare combinations in scientific discoveries and inventions tend to garner higher citation rates (*3*). Interdisciplinary research is an emblematic recombinant process (*19*); hence, the successful combination of previously disconnected ideas and resources that is fundamental to interdisciplinary research often violates expectations and leads to novel ideas with high impact (*20*). Nevertheless, evidence from grant applications shows that, when faced with new ideas, expert evaluators systematically give lower scores to truly novel (*21*–*23*) or interdisciplinary (*24*) research proposals.

The highest-impact science is primarily grounded in conventional combinations of prior work, yet it simultaneously features unusual combinations

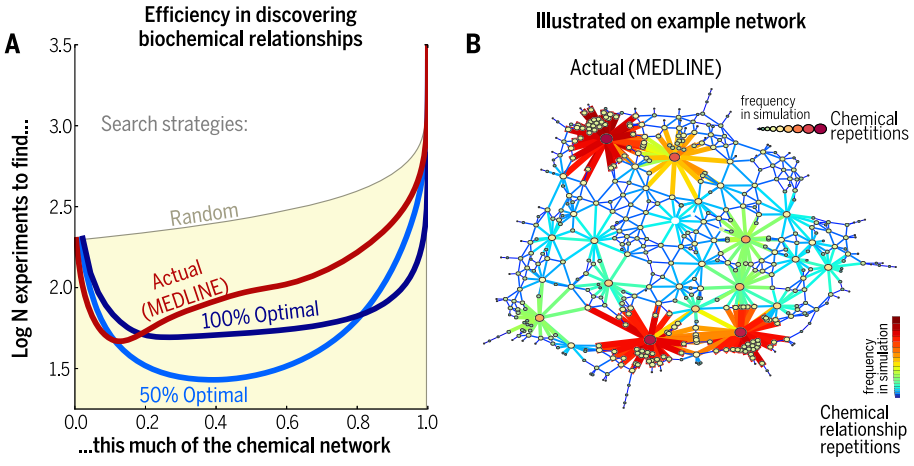


Fig. 2. Choosing experiments to accelerate collective discovery. (**A**) The average efficiency rate for global strategies to discover new, publishable chemical relationships, estimated from all MEDLINE-indexed articles published in 2010. This model does not take into account differences in the difficulty or expense of particular experiments. The efficiency of a global scientific strategy is expressed by the average number of experiments performed (vertical axis) relative to the number of new, published biochemical relationships (horizontal axis), which correspond to new connections in the published network of biochemicals co-occurring in MEDLINE-indexed articles. Compared strategies include randomly choosing pairs of biochemicals, the global ("actual") strategy inferred from all scientists publishing MEDLINE articles, and optimal strategies for discovering 50 and 100% of the network. Lower values on the vertical axis indicate more efficient strategies, showing that the actual strategy of science is suboptimal for discovering what has been published. The actual strategy is best for uncovering 13% of the chemical network, and the 50% optimal strategy is most efficient for discovering 50% of it, but neither are as good as the 100% optimal strategy for revealing the whole network. (**B**) The actual, estimated search process illustrated on a hypothetical network of chemical relationships, averaged from 500 simulated runs of that strategy. The strategy swarms around a few "important," highly connected chemicals, whereas optimal strategies are much more even and less likely to "follow the crowd" in their search across the space of scientific possibilities. [Adapted from (*15*)]

(25–27). Papers of this type are twice as likely to receive high citations (26). In other words, a balanced mixture of new and established elements is the safest path toward successful reception of scientific advances.

Career dynamics

Individual academic careers unfold in the context of a vast market for knowledge production and consumption (28). Consequently, scientific careers have been examined not only in terms of individual incentives and marginal productivity (i.e., relative gain versus effort) (29), but also institutional incentives (30, 31) and competition (32). This requires combining large repositories of high-resolution individual, geographic, and temporal metadata (33) to construct representations of career trajectories that can be analyzed from different perspectives. For example, one study finds that funding schemes that are tolerant of early failure, which reward long-term success, are more likely to generate high-impact publications than grants subject to short review cycles (31). Interacting systems with competing time scales are a classic problem in complex systems science. The multifaceted nature of science is motivation for generative models that highlight unintended consequences of policies. For example, models of career growth show that nontenure (short-term) contracts are responsible for productivity fluctuations, which often result in a sudden career death (29).

Gender inequality in science remains prevalent and problematic (34). Women have fewer publications (35–37) and collaborators (38) and less funding (39), and they are penalized in hiring decisions when compared with equally qualified men (40). The causes of these gaps are still unclear. Intrinsic differences in productivity rates and career length can explain the differences in collaboration patterns (38) and hiring rates (35) between male and female scientists. On the other hand, experimental evidence shows that biases against women occur at very early career stages. When gender was randomly assigned among the curricula vitae of a pool of applicants, the hiring committee systematically penalized female candidates (40). Most studies so far have focused on relatively small samples. Improvements in compiling large-scale data sets on scientific careers, which leverage information from different sources (e.g., publication records, grant applications, and awards), will help us gain deeper insight into the causes of inequality and motivate models that can inform policy solutions.

Scientists' mobility is another important factor offering diverse career opportunities. Most mobility studies have focused on quantifying the brain drain and gain of a country or a region (41, 42), especially after policy changes. Research on individual mobility and its career effect remains scant, however, primarily owing to the difficulty of obtaining longitudinal information about the movements of many scientists and accounts of the reasons underlying mobility decisions. Scientists who left their country of origin outperformed scientists who did not relocate, according to their citation scores, which may be rooted in a selection bias that offers better career opportunities to better scientists (43, 44). Moreover, scientists tend to move between institutions of similar prestige (45). Nevertheless, when examining changes in impact associated with each move as quantified by citations, no systematic increase or decrease was found, not even when scientists moved to an institution of considerably higher or lower rank (46). In other words, it is not the institution that creates the impact; it is the individual researchers that make an institution.

Another potentially important career factor is reputation—and the dilemma that it poses for manuscript review, proposal evaluation, and promotion decisions. The reputation of paper authors, measured by the total citations of their previous output, markedly boosts the number of citations collected by that paper in the first years after publication (47). After this initial phase, however, impact depends on the reception of the work by the scientific community. This finding, along with the work reported in (46), suggests that, for productive scientific careers, reputation is less of a critical driver for success than talent, hard work, and relevance.

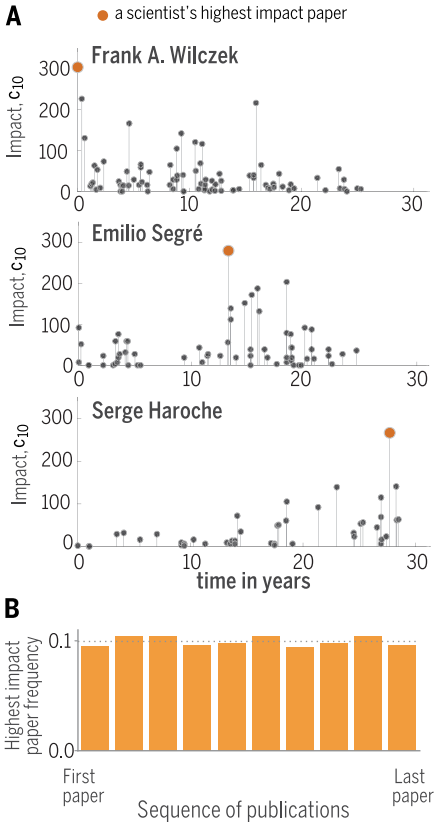


Fig. 3. Impact in scientific careers. (**A**) Publication record of three Nobel laureates in physics. The horizontal axis indicates the number of years after a laureate's first publication, each circle corresponds to a research paper, and the height of the circle represents the paper's impact, quantified by c_{10}, the number of citations after 10 years. The highest-impact paper of a laureate is denoted with an orange circle. (**B**) Histogram of the occurrence of the highest-impact paper in a scientist's sequence of publications, calculated for 10,000 scientists. The flatness of the histogram indicates that the highest-impact work can be, with the same probability, anywhere in the sequence of papers published by a scientist (49).

A policy-relevant question is whether creativity and innovation depend on age or career stage. Decades of research on outstanding researchers and innovators concluded that major breakthroughs take place relatively early in a career, with a median age of 35 (48). In contrast, recent work shows that this well-documented propensity of early-career discoveries is fully explained by productivity, which is high in the early stages of a scientist's career and drops later (49). In other words, there are no age patterns in innovation: A scholar's most cited paper can be any of his or her papers, independently of the age or career stage when it is published (Fig. 3). A stochastic model of impact evolution also indicates that breakthroughs result from a combination of the ability of a scientist and the luck of picking a problem with high potential (49).

Team science

During past decades, reliance on teamwork has increased, representing a fundamental shift in the way that science is done. A study of the authorship of 19.9 million research articles and 2.1 million patents reveals a nearly universal shift toward teams in all branches of science (50) (Fig. 4). For example, in 1955, science and engineering teams authored about the same number of papers as single authors. Yet by 2013, the fraction of team-authored papers increased to 90% (51).

Nowadays, a team-authored paper in science and engineering is 6.3 times more likely to receive 1000 citations or more than a solo-authored paper, a difference that cannot be explained by self-citations (50, 52). One possible reason is a team's ability to come up with more novel combinations of ideas (26) or to produce resources that are later used by others (e.g., genomics). Measurements show that teams are 38% more likely than solo authors to insert novel combinations into familiar knowledge domains, supporting the premise that teams can bring together different specialties, effectively combining knowledge to prompt scientific breakthroughs. Having more collaborations means greater visibility through a larger number of coauthors, who will likely introduce the work to their networks, an enhanced impact that may partially compensate for the fact that credit within a team must be shared with many colleagues (29).

Work from large teams garners, on average, more citations across a wide variety of domains. Research suggests that small teams tend to disrupt science and technology with new ideas and opportunities, whereas large teams develop existing ones (53). Thus, it may be important to

fund and foster teams of all sizes to temper the bureaucratization of science (28).

Teams are growing in size, increasing by an average of 17% per decade (50, 54), a trend underlying a fundamental change in team compositions. Scientific teams include both small, stable "core" teams and large, dynamically changing extended teams (55). The increasing team size in most fields is driven by faster expansion of extended teams, which begin as small core teams but subsequently attract new members through a process of cumulative advantage anchored by productivity. Size is a crucial determinant of team survival strategies: Small teams survive longer if they maintain a stable core, but larger teams persist longer if they manifest a mechanism for membership turnover (56).

As science has accelerated and grown increasingly complex, the instruments required to expand the frontier of knowledge have increased in scale and precision. The tools of the trade become unaffordable to most individual investigators, but also to most institutions. Collaboration has been a critical solution, pooling resources to scientific advantage. The Large Hadron Collider at CERN, the world's largest and most powerful particle collider, would have been unthinkable without collaboration, requiring more than 10,000 scientists and engineers from more than 100 countries. There is, however, a trade-off with increasing size that affects the value and risk associated with "big science" (2). Although it may be possible to solve larger problems, the burden of reproducibility may require duplicating initial efforts, which may not be practically or economically feasible.

Collaborators can have a large effect on scientific careers. According to recent studies (57, 58), scientists who lose their star collaborators experience a substantial drop in their productivity, especially if the lost collaborator was a regular coauthor. Publications involving extremely strong collaborators gain 17% more citations on average, pointing to the value of career partnership (59).

Given the increasing number of authors on the average research paper, who should and does gain the most credit? The canonical theory of credit (mis)allocation in science is the Matthew effect (60), in which scientists of higher statuses involved in joint work receive outsized credit for their contributions. Properly allocating individual credit for a collaborative work is difficult because we cannot easily distinguish individual contributions (61). It is possible, however, to inspect the co-citation patterns of the coauthors' publications to determine the fraction of credit that the community assigns to each coauthor in a publication (62).

Citation dynamics

Scholarly citation remains the dominant measurable unit of credit in science. Given the reliance of most impact metrics on citations (63–66), the dynamics of citation accumulation have been scrutinized by generations of scholars. From foundational work by Price (67), we know that the distribution of citations for scientific papers is highly skewed: Many papers are never cited, but seminal papers can accumulate 10,000 or more citations. This uneven citation distribution is a robust, emergent property of the dynamics of science, and it holds when papers are grouped by institution (68). If the number of citations of a paper is divided by the average number of citations collected by papers in the same discipline and year, the distribution of the resulting score is essentially indistinguishable for all disciplines (69, 70) (Fig. 5A). This means that we can compare the impact of papers published in different disciplines by looking at their relative citation values. For example, a paper in mathematics collecting 100 citations represents a higher disciplinary impact than a paper in microbiology with 300 citations.

The tail of the citation distribution, capturing the number of high-impact papers, sheds light on the mechanisms that drive the accumulation of citations. Recent analyses show that it follows a power law (71–73). Power-law tails can be generated through a cumulative advantage process (74), known as preferential attachment in network science (75), suggesting that the probability of citing a paper grows with the number of citations that it has already collected. Such a model can be augmented with other characteristic features of citation dynamics, such as the obsolescence of knowledge, decreasing the citation probability with the age of the paper (76–79), and a fitness parameter, unique to each paper, capturing the appeal of the work to the scientific community (77, 78). Only a tiny fraction of papers deviate from the pattern described by such a model—some of which are called "sleeping beauties," because they receive very little notice for decades after publication and then suddenly receive a burst of attention and citations (80, 81).

The generative mechanisms described above can be used to predict the citation dynamics of individual papers. One predictive model (77) assumes that the citation probability of a paper depends on the number of previous citations, an obsolescence factor, and a fitness parameter (Fig. 5, B and C). For a given paper, one can estimate the three model parameters by fitting the model to the initial portion of the citation history of the paper. The long-term impact of the work can be extrapolated (77). Other studies have identified predictors of the citation impact of individual papers (82), such as journal impact factor (72). It has been suggested that the future h-index (83) of a scientist can be accurately predicted (84), although the predictive power is reduced when

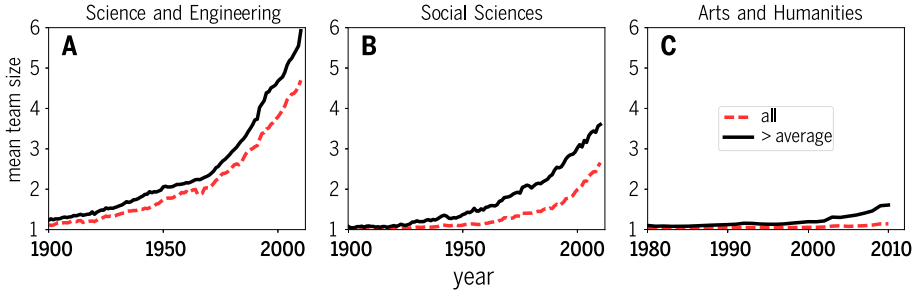


Fig. 4. Size and impact of teams. Mean team size has been steadily growing over the past century. The red dashed curves represent the mean number of coauthors over all papers; the black curves consider just those papers receiving more citations than the average for the field. Black curves are systematically above the dashed red ones, meaning that high-impact work is more likely to be produced by large teams than by small ones. Each panel corresponds to one of the three main disciplinary groups of papers indexed in the WoS: (**A**) science and engineering, (**B**) social sciences, and (**C**) arts and humanities.

Box 1. Lessons from SciSci.

1. Innovation and tradition: Left bare, truly innovative and highly interdisciplinary ideas may not reach maximum scientific impact. To enhance their impact, novel ideas should be placed in the context of established knowledge (26).
2. Persistence: A scientist is never too old to make a major discovery, as long as he or she stays productive (49).
3. Collaboration: Research is shifting to teams, so engaging in collaboration is beneficial. Works by small teams tend to be more disruptive, whereas those by big teams tend to have more impact (4, 50, 53).
4. Credit: Most credit will go to the coauthors with the most consistent track record in the domain of the publication (62).
5. Funding: Although review panels acknowledge innovation, they ultimately tend to discount it. Funding agencies should ask reviewers to assess innovation, not only expected success (24).

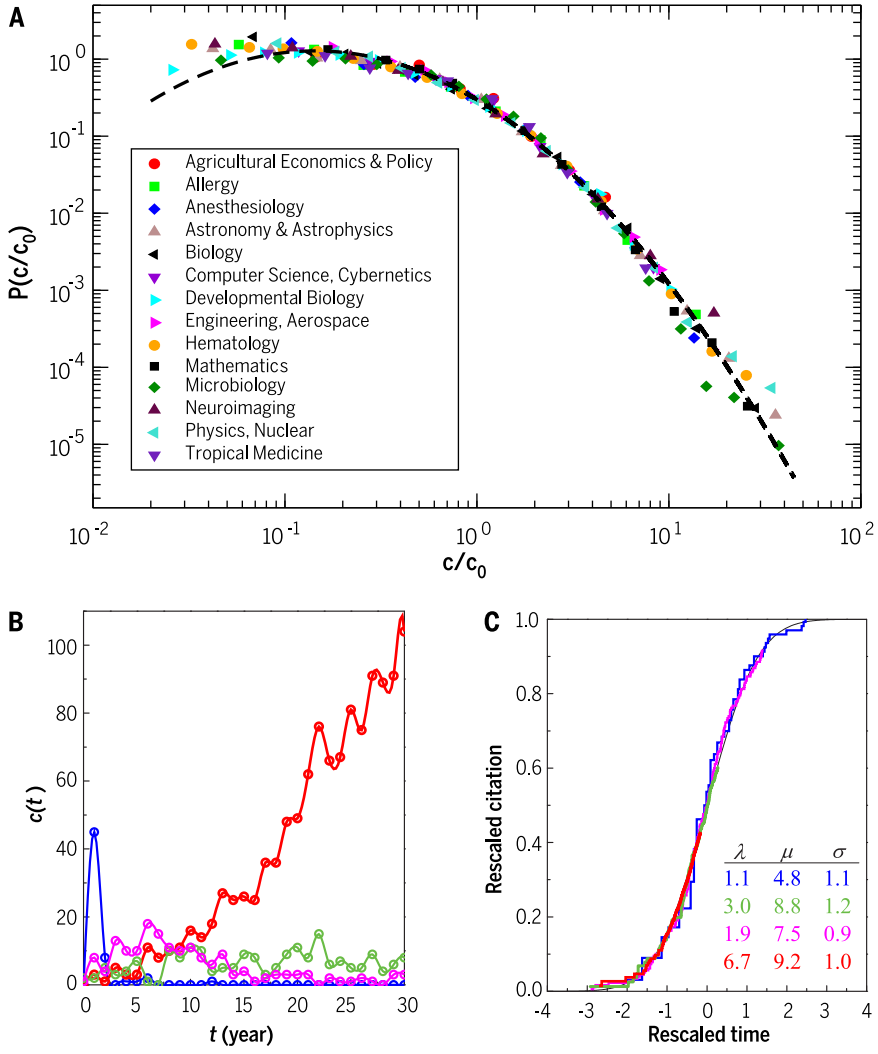


Fig. 5. Universality in citation dynamics. (**A**) The citation distributions of papers published in the same discipline and year lie on the same curve for most disciplines, if the raw number of citations c of each paper is divided by the average number of citations c_0 over all papers in that discipline and year. The dashed line is a lognormal fit. [Adapted from (*69*)] (**B**) Citation history of four papers published in *Physical Review* in 1964, selected for their distinct dynamics, displaying a "jump-decay" pattern (blue), experiencing a delayed peak (magenta), attracting a constant number of citations over time (green), or acquiring an increasing number of citations each year (red). (**C**) Citations of an individual paper are determined by three parameters: fitness λ_i, immediacy μ_i, and longevity σ_i. By rescaling the citation history of each paper in (B) by the appropriate (λ, μ, σ) parameters, the four papers collapse onto a single universal function, which is the same for all disciplines. [Adapted from (*77*)]

accounting for the scientist's career stage and the cumulative, nondecreasing nature of the h-index (*85*). Eliminating inconsistencies in the use of quantitative evaluation metrics in science is crucial and highlights the importance of understanding the generating mechanisms behind commonly used statistics.

Outlook

Despite the discovery of universals across science, substantial disciplinary differences in culture, habits, and preferences make some cross-domain insights difficult to appreciate within particular fields and associated policies challenging to implement. The differences among the questions, data, and skills required by each discipline suggest that we may gain further insights from domain-specific SciSci studies that model and predict opportunities adapted to the needs of each field. For young scientists, the results of SciSci offer actionable insights about past patterns, helping guide future inquiry within their disciplines (Box 1).

The contribution of SciSci is a detailed understanding of the relational structure between scientists, institutions, and ideas, a crucial starting point that facilitates the identification of fundamental generating processes. Together, these data-driven efforts complement contributions from related research domains such as the economics (*30*) and sociology of science (*60*, *86*). Causal estimation is a prime example, in which econometric matching techniques demand and leverage comprehensive data sources in the effort to simulate counterfactual scenarios (*31*, *42*). Assessing causality is one of the most needed future developments in SciSci: Many descriptive studies reveal strong associations between structure and outcomes, but the extent to which a specific structure "causes" an outcome remains unexplored. Engaging in tighter partnerships with experimentalists, SciSci will be able to better identify associations discovered from models and large-scale data that have causal force to enrich their policy relevance. But experimenting on science may be the biggest challenge SciSci has yet to face. Running randomized, controlled trials that can alter outcomes for individuals or institutions of science, which are mostly supported by tax dollars, is bound to elicit criticisms and pushback (*87*). Hence, we expect quasi-experimental approaches to prevail in SciSci investigations in the near future.

Most SciSci research focuses on publications as primary data sources, implying that insights and findings are limited to ideas successful enough to merit publication in the first place. Yet most scientific attempts fail, sometimes spectacularly. Given that scientists fail more often than they succeed, knowing when, why, and how an idea fails is essential in our attempts to understand and improve science. Such studies could provide meaningful guidance regarding the reproducibility crisis and help us account for the file drawer problem. They could also substantially further our understanding of human imagination by revealing the total pipeline of creative activity.

Science often behaves like an economic system with a one-dimensional "currency" of citation counts. This creates a hierarchical system, in which the "rich-get-richer" dynamics suppress the spread of new ideas, particularly those from junior scientists and those who do not fit within the paradigms supported by specific fields. Science can be improved by broadening the number and range of performance indicators. The development of alternative metrics covering web (*88*) and social media (*89*) activity and societal impact (*90*) is critical in this regard. Other measurable dimensions include the information (e.g., data) that scientists share with competitors (*91*), the help that they offer to their peers (*92*), and their reliability as reviewers of their peers' works (*93*). But with a profusion of metrics, more work is needed to understand what each of them does and does not capture to ensure meaningful interpretation and avoid misuse. SciSci can make an essential contribution by providing models that offer a deeper understanding of the mechanisms that govern performance indicators in science. For instance, models of the empirical patterns observed when alternative indicators (e.g., distributions of paper downloads) are used will enable us to explore their relationship with citation-based metrics (*94*) and to recognize manipulations.

The integration of citation-based metrics with alternative indicators will promote pluralism and enable new dimensions of productive specialization, in which scientists can be successful in different ways. Science is an ecosystem that requires not only publications, but also communicators, teachers, and detail-oriented experts. We need individuals who can ask novel, field-altering questions, as well as those who can answer them. It would benefit science if curiosity, creativity, and intellectual exchange—particularly regarding the societal implications and applications of science and technology—are better appreciated and incentivized in the future. A more pluralistic approach could reduce duplication and make science flourish for society (*95*).

An issue that SciSci seeks to address is the allocation of science funding. The current peer review system is subject to biases and inconsistencies (*96*). Several alternatives have been proposed, such as the random distribution of funding (*97*), person-directed funding that does not involve proposal preparation and review (*31*), opening the proposal review process to the entire online population (*98*), removing human reviewers altogether by allocating funds through a performance measure (*99*), and scientist crowdfunding (*100*).

A critical area of future research for SciSci concerns the integration of machine learning and artificial intelligence in a way that involves machines and minds working together. These new tools portend far-reaching implications for science because machines might broaden a scientist's perspective more than human collaborators. For instance, the self-driving vehicle is the result of a successful combination of known driving habits and information that was outside of human awareness, provided by sophisticated machine-learning techniques. Mind-machine partnerships have improved evidence-based decision-making in a wide range of health, economic, social, legal, and business problems (*101*–*103*). How can science be improved with mind-machine partnerships, and what arrangements are most productive? These questions promise to help us understand the science of the future.

REFERENCES AND NOTES

1. E. Garfield, Citation indexes for science; a new dimension in documentation through association of ideas. *Science* **122**, 108–111 (1955). doi: 10.1126/science.122.3159.108; pmid: 14385826
2. D. J. S. Price, *Little Science, Big Science* (Columbia Univ. Press, 1963).
3. J. G. Foster, A. Rzhetsky, J. A. Evans, Tradition and innovation in scientists' research strategies. *Am. Sociol. Rev.* **80**, 875–908 (2015). doi: 10.1177/0003122415601618
4. S. Milojević, Quantifying the cognitive extent of science. *J. Informetr.* **9**, 962–973 (2015). doi: 10.1016/j.joi.2015.10.005
5. T. Kuhn, M. Perc, D. Helbing, Inheritance patterns in citation networks reveal scientific memes. *Phys. Rev. X* **4**, 041036 (2014). doi: 10.1103/PhysRevX.4.041036
6. R. Klavans, K. W. Boyack, Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge? *J. Assoc. Inf. Sci. Technol.* **68**, 984–998 (2016). doi: 10.1002/asi.23734
7. U. Shwed, P. S. Bearman, The temporal structure of scientific consensus formation. *Am. Sociol. Rev.* **75**, 817–840 (2010). doi: 10.1177/0003122410388488; pmid: 21886269
8. J. Bruggeman, V. A. Traag, J. Uitermark, Detecting communities through network data. *Am. Sociol. Rev.* **77**, 1050–1063 (2012). doi: 10.1177/0003122412463574
9. F. Shi, J. G. Foster, J. A. Evans, Weaving the fabric of science: Dynamic network models of science's unfolding structure. *Soc. Networks* **43**, 73–85 (2015). doi: 10.1016/j.socnet.2015.02.006
10. L. M. A. Bettencourt, D. I. Kaiser, J. Kaur, Scientific discovery and topological transitions in collaboration networks. *J. Informetr.* **3**, 210–221 (2009). doi: 10.1016/j.joi.2009.03.001
11. X. Sun, J. Kaur, S. Milojević, A. Flammini, F. Menczer, Social dynamics of science. *Sci. Rep.* **3**, 1069 (2013). doi: 10.1038/srep01069; pmid: 23323212
12. T. S. Kuhn, *The Essential Tension: Selected Studies in Scientific Tradition and Change* (Univ. of Chicago Press, 1977).
13. P. Bourdieu, The specificity of the scientific field and the social conditions of the progress of reasons. *Soc. Sci. Inf. (Paris)* **14**, 19–47 (1975). doi: 10.1177/053901847501400602
14. T. Jia, D. Wang, B. K. Szymanski, Quantifying patterns of research-interest evolution. *Nat. Hum. Behav.* **1**, 0078 (2017). doi: 10.1038/s41562-017-0078
15. A. Rzhetsky, J. G. Foster, I. T. Foster, J. A. Evans, Choosing experiments to accelerate collective discovery. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 14569–14574 (2015). doi: 10.1073/pnas.1509757112; pmid: 26554009
16. R. Rosenthal, The file drawer problem and tolerance for null results. *Psychol. Bull.* **86**, 638–641 (1979). doi: 10.1037/0033-2909.86.3.638
17. S. B. Nissen, T. Magidson, K. Gross, C. T. Bergstrom, Publication bias and the canonization of false facts. *eLife* **5**, e21451 (2016). doi: 10.7554/eLife.21451; pmid: 27995896
18. L. Yao, Y. Li, S. Ghosh, J. A. Evans, A. Rzhetsky, Health ROI as a measure of misalignment of biomedical needs and resources. *Nat. Biotechnol.* **33**, 807–811 (2015). doi: 10.1038/nbt.3276; pmid: 26252133
19. C. S. Wagner *et al.*, Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *J. Informetr.* **5**, 14–26 (2011). doi: 10.1016/j.joi.2010.06.004
20. V. Larivière, S. Haustein, K. Börner, Long-distance interdisciplinarity leads to higher scientific impact. *PLOS ONE* **10**, e0122565 (2015). doi: 10.1371/journal.pone.0122565; pmid: 25822658
21. K. J. Boudreau, E. C. Guinan, K. R. Lakhani, C. Riedl, Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Manage. Sci.* **62**, 2765–2783 (2016). doi: 10.1287/mnsc.2015.2285; pmid: 27746512
22. E. Leahey, J. Moody, Sociological innovation through subfield integration. *Soc. Currents* **1**, 228–256 (2014). doi: 10.1177/2329496514540131
23. A. Yegros-Yegros, I. Rafols, P. D'Este, Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity. *PLOS ONE* **10**, e0135095 (2015). doi: 10.1371/journal.pone.0135095; pmid: 26266805
24. L. Bromham, R. Dinnage, X. Hua, Interdisciplinary research has consistently lower funding success. *Nature* **534**, 684–687 (2016). doi: 10.1038/nature18315; pmid: 27357795
25. D. Kim, D. B. Cerigo, H. Jeong, H. Youn, Technological novelty profile and inventions future impact. *EPJ Data Sci.* **5**, 8 (2016). doi: 10.1140/epjds/s13688-016-0069-1
26. B. Uzzi, S. Mukherjee, M. Stringer, B. Jones, Atypical combinations and scientific impact. *Science* **342**, 468–472 (2013). doi: 10.1126/science.1240474; pmid: 24159044
27. J. Wang, R. Veugelers, P. Stephan, "Bias against novelty in science: A cautionary tale for users of bibliometric indicators" (NBER Working Paper No. 22180, National Bureau of Economic Research, 2016).
28. J. P. Walsh, Y.-N. Lee, The bureaucratization of science. *Res. Policy* **44**, 1584–1600 (2015). doi: 10.1016/j.respol.2015.04.010
29. A. M. Petersen, M. Riccaboni, H. E. Stanley, F. Pammolli, Persistence and uncertainty in the academic career. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5213–5218 (2012). doi: 10.1073/pnas.1121429109; pmid: 22431620
30. P. E. Stephan, *How Economics Shapes Science* (Harvard Univ. Press, 2012).
31. P. Azoulay, J. S. Graff Zivin, G. Manso, Incentives and creativity: Evidence from the academic life sciences. *Rand J. Econ.* **42**, 527–554 (2011). doi: 10.1111/j.1756-2171.2011.00140.x
32. R. Freeman, E. Weinstein, E. Marincola, J. Rosenbaum, F. Solomon, Competition and careers in biosciences. *Science* **294**, 2293–2294 (2001). doi: 10.1126/science.1067477; pmid: 11743184
33. J. A. Evans, J. G. Foster, Metaknowledge. *Science* **331**, 721–725 (2011). doi: 10.1126/science.1201765; pmid: 21311014
34. V. Larivière, C. Ni, Y. Gingras, B. Cronin, C. R. Sugimoto, Bibliometrics: Global gender disparities in science. *Nature* **504**, 211–213 (2013). doi: 10.1038/504211a; pmid: 24350369
35. S. F. Way, D. B. Larremore, A. Clauset, in *Proceedings of the 25th International Conference on World Wide Web (WWW '16)* (ACM, 2016), pp. 1169–1179.
36. J. Duch *et al.*, The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PLOS ONE* **7**, e51332 (2012). doi: 10.1371/journal.pone.0051332; pmid: 23251502
37. J. D. West, J. Jacquet, M. M. King, S. J. Correll, C. T. Bergstrom, The role of gender in scholarly authorship. *PLOS ONE* **8**, e66212 (2013). doi: 10.1371/journal.pone.0066212; pmid: 23894278
38. X. H. T. Zeng *et al.*, Differences in collaboration patterns across discipline, career stage, and gender. *PLOS Biol.* **14**, e1002573 (2016). doi: 10.1371/journal.pbio.1002573; pmid: 27814355
39. T. J. Ley, B. H. Hamilton, The gender gap in NIH grant applications. *Science* **322**, 1472–1474 (2008). doi: 10.1126/science.1165878; pmid: 19056961
40. C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham, J. Handelsman, Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16474–16479 (2012). doi: 10.1073/pnas.1211286109; pmid: 22988126
41. R. Van Noorden, Global mobility: Science on the move. *Nature* **490**, 326–329 (2012). doi: 10.1038/490326a; pmid: 23075963
42. O. A. Doria Arrieta, F. Pammolli, A. M. Petersen, Quantifying the negative impact of brain drain on the integration of European science. *Sci. Adv.* **3**, e1602232 (2017). doi: 10.1126/sciadv.1602232; pmid: 28439544
43. C. Franzoni, G. Scellato, P. Stephan, The mover's advantage: The superior performance of migrant scientists. *Econ. Lett.* **122**, 89–93 (2014). doi: 10.1016/j.econlet.2013.10.040
44. C. R. Sugimoto *et al.*, Scientists have most impact when they're free to move. *Nature* **550**, 29–31 (2017). doi: 10.1038/550029a; pmid: 28980663
45. A. Clauset, S. Arbesman, D. B. Larremore, Systematic inequality and hierarchy in faculty hiring networks. *Sci. Adv.* **1**, e1400005 (2015). doi: 10.1126/sciadv.1400005; pmid: 26601125
46. P. Deville *et al.*, Career on the move: Geography, stratification, and scientific impact. *Sci. Rep.* **4**, 4770 (2014). pmid: 24759743
47. A. M. Petersen *et al.*, Reputation and impact in academic careers. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15316–15321 (2014). doi: 10.1073/pnas.1323111111; pmid: 25288774
48. D. K. Simonton, Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychol. Rev.* **104**, 66–89 (1997). doi: 10.1037/0033-295X.104.1.66
49. R. Sinatra, D. Wang, P. Deville, C. Song, A.-L. Barabási, Quantifying the evolution of individual scientific impact. *Science* **354**, aaf5239 (2016). doi: 10.1126/science.aaf5239; pmid: 27811240
50. S. Wuchty, B. F. Jones, B. Uzzi, The increasing dominance of teams in production of knowledge. *Science* **316**, 1036–1039 (2007). doi: 10.1126/science.1136099; pmid: 17431139
51. N. J. Cooke, M. L. Hilton, Eds., *Enhancing the Effectiveness of Team Science* (National Academies Press, 2015).
52. V. Larivière, Y. Gingras, C. R. Sugimoto, A. Tsou, Team size matters: Collaboration and scientific impact since 1900. *J. Assoc. Inf. Sci. Technol.* **66**, 1323–1332 (2015). doi: 10.1002/asi.23266
53. L. Wu, D. Wang, J. A. Evans, Large teams have developed science and technology; small teams have disrupted it. arXiv:1709.02445 [physics.soc-ph] (7 September 2017).
54. B. F. Jones, The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?

Rev. Econ. Stud. **76**, 283–317 (2009). doi: 10.1111/j.1467-937X.2008.00531.x
55. S. Milojević, Principles of scientific research team formation and evolution. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3984–3989 (2014). doi: 10.1073/pnas.1309723111; pmid: 24591626
56. G. Palla, A.-L. Barabási, T. Vicsek, Quantifying social group evolution. *Nature* **446**, 664–667 (2007). doi: 10.1038/nature05670; pmid: 17410175
57. G. J. Borjas, K. B. Doran, Which peers matter? The relative impacts of collaborators, colleagues, and competitors. *Rev. Econ. Stat.* **97**, 1104–1117 (2015). doi: 10.1162/REST_a_00472
58. P. Azoulay, J. G. Zivin, J. Wang, Superstar extinction. *Q. J. Econ.* **125**, 549–589 (2010). doi: 10.1162/qjec.2010.125.2.549
59. A. M. Petersen, Quantifying the impact of weak, strong, and super ties in scientific careers. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E4671–E4680 (2015). doi: 10.1073/pnas.1501444112; pmid: 26261301
60. R. K. Merton, The Matthew effect in science. *Science* **159**, 56–63 (1968). doi: 10.1126/science.159.3810.56
61. L. Allen, J. Scott, A. Brand, M. Hlava, M. Altman, Publishing: Credit where credit is due. *Nature* **508**, 312–313 (2014). doi: 10.1038/508312a; pmid: 24745070
62. H.-W. Shen, A.-L. Barabási, Collective credit allocation in science. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 12325–12330 (2014). doi: 10.1073/pnas.1401992111; pmid: 25114238
63. L. Waltman, A review of the literature on citation impact indicators. *J. Informetr.* **10**, 365–391 (2016). doi: 10.1016/j.joi.2016.02.007
64. J. E. Hirsch, An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16569–16572 (2005). doi: 10.1073/pnas.0507655102; pmid: 16275915
65. H. F. Moed, *Citation Analysis in Research Evaluation* (Springer, 2010).
66. E. Garfield, Citation analysis as a tool in journal evaluation. *Science* **178**, 471–479 (1972). doi: 10.1126/science.178.4060.471; pmid: 5079701
67. D. J. de Solla Price, Networks of scientific papers. *Science* **149**, 510–515 (1965). doi: 10.1126/science.149.3683.510; pmid: 14325149
68. Q. Zhang, N. Perra, B. Gonçalves, F. Ciulla, A. Vespignani, Characterizing scientific production and consumption in physics. *Sci. Rep.* **3**, 1640 (2013). doi: 10.1038/srep01640; pmid: 23571320
69. F. Radicchi, S. Fortunato, C. Castellano, Universality of citation distributions: Toward an objective measure of scientific impact. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17268–17272 (2008). doi: 10.1073/pnas.0806977105; pmid: 18978030
70. L. Waltman, N. J. van Eck, A. F. J. van Raan, Universality of citation distributions revisited. *J. Assoc. Inf. Sci. Technol.* **63**, 72–77 (2012). doi: 10.1002/asi.21671
71. M. Golosovsky, S. Solomon, Runaway events dominate the heavy tail of citation distributions. *Eur. Phys. J. Spec. Top.* **205**, 303–311 (2012). doi: 10.1140/epjst/e2012-01576-4
72. C. Stegehuis, N. Litvak, L. Waltman, Predicting the long-term citation impact of recent publications. *J. Informetr.* **9**, 642–657 (2015). doi: 10.1016/j.joi.2015.06.005
73. M. Thelwall, The discretised lognormal and hooked power law distributions for complete citation data: Best options for modelling and regression. *J. Informetr.* **10**, 336–346 (2016). doi: 10.1016/j.joi.2015.12.007
74. D. de Solla Price, A general theory of bibliometric and other cumulative advantage processes. *J. Am. Soc. Inf. Sci.* **27**, 292–306 (1976). doi: 10.1002/asi.4630270505
75. A.-L. Barabási, R. Albert, Emergence of scaling in random networks. *Science* **286**, 509–512 (1999). doi: 10.1126/science.286.5439.509; pmid: 10521342
76. P. D. B. Parolo *et al.*, Attention decay in science. *J. Informetr.* **9**, 734–745 (2015). doi: 10.1016/j.joi.2015.07.006
77. D. Wang, C. Song, A.-L. Barabási, Quantifying long-term scientific impact. *Science* **342**, 127–132 (2013). doi: 10.1126/science.1237825; pmid: 24092745
78. Y.-H. Eom, S. Fortunato, Characterizing and modeling citation dynamics. *PLOS ONE* **6**, e24926 (2011). doi: 10.1371/journal.pone.0024926; pmid: 21966387
79. M. Golosovsky, S. Solomon, Stochastic dynamical model of a growing citation network based on a self-exciting point process. *Phys. Rev. Lett.* **109**, 098701 (2012). doi: 10.1103/PhysRevLett.109.098701; pmid: 23002894
80. A. F. J. van Raan, Sleeping Beauties in science. *Scientometrics* **59**, 467–472 (2004). doi: 10.1023/B:SCIE.0000018543.82441.f1
81. Q. Ke, E. Ferrara, F. Radicchi, A. Flammini, Defining and identifying Sleeping Beauties in science. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7426–7431 (2015). doi: 10.1073/pnas.1424329112; pmid: 26015563
82. I. Tahamtan, A. Safipour Afshar, K. Ahamdzadeh, Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics* **107**, 1195–1225 (2016). doi: 10.1007/s11192-016-1889-2
83. J. E. Hirsch, Does the *h* index have predictive power? *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19193–19198 (2007). doi: 10.1073/pnas.0707962104; pmid: 18040045
84. D. E. Acuna, S. Allesina, K. P. Kording, Future impact: Predicting scientific success. *Nature* **489**, 201–202 (2012). doi: 10.1038/489201a; pmid: 22972278
85. O. Penner, R. K. Pan, A. M. Petersen, K. Kaski, S. Fortunato, On the predictability of future impact in science. *Sci. Rep.* **3**, 3052 (2013). doi: 10.1038/srep03052; pmid: 24165898
86. J. R. Cole, H. Zuckerman, in *The Idea of Social Structure: Papers in Honor of Robert K. Merton*, L. A. Coser, Ed. (Harcourt Brace Jovanovich, 1975), pp. 139–174.
87. P. Azoulay, Research efficiency: Turn the scientific method on ourselves. *Nature* **484**, 31–32 (2012). doi: 10.1038/484031a; pmid: 22481340
88. M. Thelwall, K. Kousha, Web indicators for research evaluation. Part 1: Citations and links to academic articles from the Web. *Prof. Inf.* **24**, 587–606 (2015). doi: 10.3145/epi.2015.sep.08
89. M. Thelwall, K. Kousha, Web indicators for research evaluation. Part 2: Social media metrics. *Prof. Inf.* **24**, 607–620 (2015). doi: 10.3145/epi.2015.sep.09
90. L. Bornmann, What is societal impact of research and how can it be assessed? A literature survey. *Adv. Inf. Sci.* **64**, 217–233 (2013).
91. C. Haeussler, L. Jiang, J. Thursby, M. Thursby, Specific and general information sharing among competing academic researchers. *Res. Policy* **43**, 465–475 (2014). doi: 10.1016/j.respol.2013.08.017
92. A. Oettl, Sociology: Honour the helpful. *Nature* **489**, 496–497 (2012). doi: 10.1038/489496a; pmid: 23018949
93. S. Ravindran, "Getting credit for peer review," *Science*, 8 February 2016; www.sciencemag.org/careers/2016/02/getting-credit-peer-review.
94. R. Costas, Z. Zahedi, P. Wouters, Do "altmetrics" correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *J. Assoc. Inf. Sci. Technol.* **66**, 2003–2019 (2015). doi: 10.1002/asi.23309
95. A. Clauset, D. B. Larremore, R. Sinatra, Data-driven predictions in the science of science. *Science* **355**, 477–480 (2017). doi: 10.1126/science.aal4217
96. S. Wessely, Peer review of grant applications: What do we know? *Lancet* **352**, 301–305 (1998). doi: 10.1016/S0140-6736(97)11129-1; pmid: 9690424
97. N. Geard, J. Noble, paper presented at the 3rd World Congress on Social Simulation, Kassel, Germany, 6 to 9 September 2010.
98. Calm in a crisis. *Nature* **468**, 1002 (2010). doi: 10.1038/4681002a; pmid: 21170024
99. R. Roy, Funding science: The real defects of peer review and an alternative to it. *Sci. Technol. Human Values* **10**, 73–81 (1985). doi: 10.1177/016224398501000309
100. J. Bollen, D. Crandall, D. Junk, Y. Ding, K. Börner, An efficient system to fund science: From proposal review to peer-to-peer distributions. *Scientometrics* **110**, 521–528 (2017). doi: 10.1007/s11192-016-2110-3
101. M. S. Kohn *et al.*, IBM's health analytics and clinical decision support. *Yearb. Med. Inform.* **9**, 154–162 (2014). doi: 10.15265/IY-2014-0002; pmid: 25123736
102. J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, S. Mullainathan, "Human decisions and machine predictions" (National Bureau of Economic Research, 2017).
103. B. Liu, R. Govindan, B. Uzzi, Do emotions expressed online correlate with actual changes in decision-making?: The case of stock day traders. *PLOS ONE* **11**, e0144945 (2016). doi: 10.1371/journal.pone.0144945; pmid: 26765539

ACKNOWLEDGMENTS

This work was supported by Air Force Office of Scientific Research grants FA9550-15-1-0077 (A.-L.B., R.S., and A.V.), FA9550-15-1-0364 (A.-L.B. and R.S.), FA9550-15-1-0162 (J.A.E. and D.W.), and FA9550-17-1-0089 (D.W.); National Science Foundation grants NCSE 1538763, EAGER 1566393, and NCN CP supplement 1553044 (K.B.) and SBE1158803 (J.A.E.); National Institutes of Health awards P01 AG039347 and U01CA198934 (K.B.) and IIS-0910664 (B.U.); Army Research Office grant W911NF-15-1-0577 and Northwestern University Institute on Complex Systems (B.U.); DARPA (Defense Advanced Research Projects Agency) Big Mechanism program grant 14145043 and the John Templeton Foundation's grant to the Metaknowledge Network (J.A.E.); Intellectual Themes Initiative "Just Data" project (R.S.); and European Commission H2020 FETPROACT-GSS CIMPLEX grant 641191 (R.S. and A.-L.B.). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of our funders.

10.1126/science.aao0185

RESEARCH ARTICLE SUMMARY

MICROBIOLOGY

Systematic discovery of antiphage defense systems in the microbial pangenome

Shany Doron,* Sarah Melamed,* Gal Ofir, Azita Leavitt, Anna Lopatina, Mai Keren, Gil Amitai, Rotem Sorek†

INTRODUCTION: Bacteria and archaea are frequently attacked by viruses (phages) and as a result have developed multiple, sophisticated lines of active defense that can collectively be referred to as the prokaryotic "immune system." Although bacterial defense against phages has been studied for decades, it was suggested that many currently unknown defense systems reside in the genomes of nonmodel bacteria and archaea and await discovery.

RATIONALE: Antiphage defense systems are known to be frequently physically clustered in microbial genomes such that, for example, genes encoding restriction enzymes commonly reside in the vicinity of genes encoding other phage resistance systems. The observation that defense systems are clustered in genomic "defense islands" has led to the hypothesis that genes of unknown function residing within such defense islands may also participate in antiphage defense. In this study, we aimed to comprehensively identify and experimentally verify new defense systems based on their enrichment within defense islands in an attempt to systematically map the arsenal of defense tools that are at the disposal of microbes in their fight against phages.

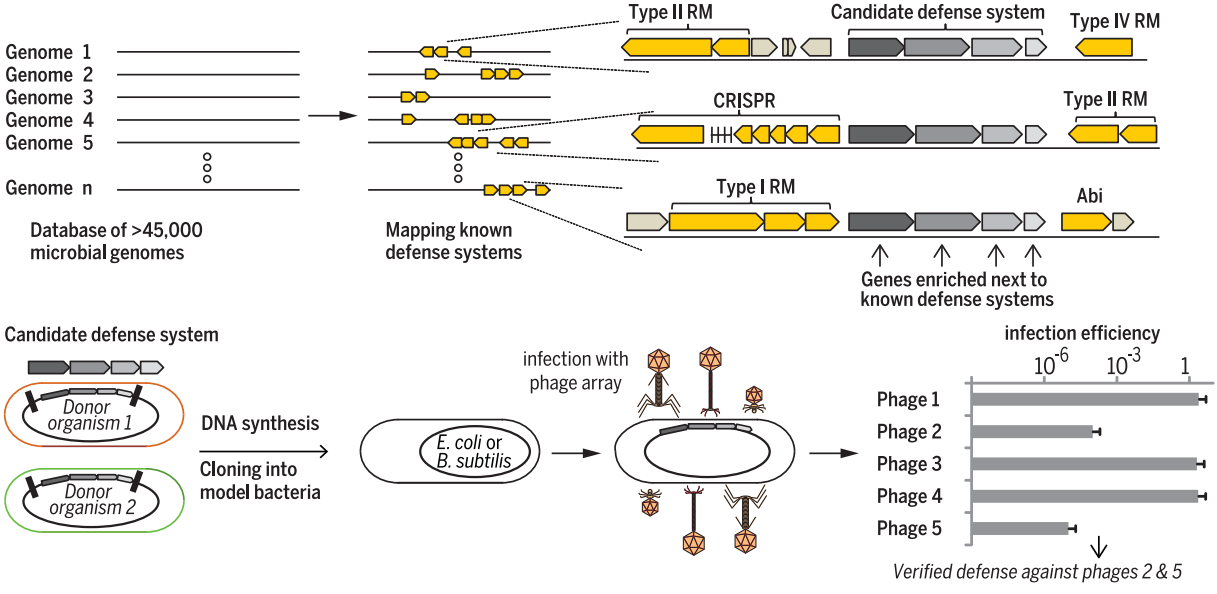
RESULTS: We searched for gene cassettes of unknown function that are enriched near known defense systems in more than 45,000 available bacterial and archaeal genome sequences. Such gene cassettes were defined as candidate defense systems and were systematically engineered into model bacteria, which were then infected by an array of phages to test for antiphage activities. This yielded the discovery of nine new families of antiphage defense systems and one additional family of antiplasmid systems that are widespread in microbes and shown to strongly protect against foreign DNA invasion. The systems discovered include ones that seem to have adopted components of the bacterial flagella and chromosome maintenance complexes and use these components for defensive capacities. Our data also show that genes with Toll-interleukin receptor (TIR) domains are involved in bacterial defense against phages, providing evidence for a common, ancient ancestry of innate immunity components shared between animals, plants, and bacteria.

CONCLUSIONS: Our study expands the known arsenal of defense systems used by prokaryotes for protection against phages, exposing tens of thousands of instances of defense systems that were so far unknown. Some of these systems appear to employ completely new mechanisms of defense against phages. In the past, the discovery and mechanistic understanding of antiphage defense systems led to the development of important biotechnological tools, as exemplified by the use of restriction enzymes and CRISPR-Cas for biotechnological and biomedical applications. One may envision that some of the systems discovered in the current study, once their mechanism is deciphered, will also be adapted into useful molecular tools in the future. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aar4120

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: rotem.sorek@weizmann.ac.il
Cite this article as S. Doron *et al.*, *Science* **359**, eaar4120 (2018). DOI: 10.1126/science.aar4120



A pipeline for systematic discovery of defense systems. Microbial genomes (more than 45,000 in the current study) are mined for genetic systems that are physically enriched next to known defense systems such as restriction-modification and CRISPR-Cas. Candidate predicted systems are cloned into model bacteria, and these bacteria are then infected by an array of phages from various families to determine whether they provide defense.

RESEARCH ARTICLE

MICROBIOLOGY

Systematic discovery of antiphage defense systems in the microbial pangenome

Shany Doron,* Sarah Melamed,* Gal Ofir, Azita Leavitt, Anna Lopatina, Mai Keren, Gil Amitai, Rotem Sorek†

The arms race between bacteria and phages led to the development of sophisticated antiphage defense systems, including CRISPR-Cas and restriction-modification systems. Evidence suggests that known and unknown defense systems are located in "defense islands" in microbial genomes. Here, we comprehensively characterized the bacterial defensive arsenal by examining gene families that are clustered next to known defense genes in prokaryotic genomes. Candidate defense systems were systematically engineered and validated in model bacteria for their antiphage activities. We report nine previously unknown antiphage systems and one antiplasmid system that are widespread in microbes and strongly protect against foreign invaders. These include systems that adopted components of the bacterial flagella and condensin complexes. Our data also suggest a common, ancient ancestry of innate immunity components shared between animals, plants, and bacteria.

Bacteria and archaea are frequently attacked by viruses (phages) and as a result have developed multiple, sophisticated lines of active defense (*1*–*3*) that can collectively be referred to as the prokaryotic "immune system." Antiphage defense strategies include restriction-modification (R-M) systems that target specific sequences on the invading phage (*4*); CRISPR-Cas, which provides acquired immunity through memorization of past phage attacks (*5*); abortive infection systems (Abi) that lead to cell death or metabolic arrest upon infection (*6*); and additional systems whose mechanism of action is not yet clear, such as BREX (*7*), prokaryotic Argonautes (pAgos) (*8*), and DISARM (*9*). Different bacteria encode different sets of defense systems: CRISPR-Cas systems are found in about 40% of all sequenced bacteria (*10*, *11*), R-M systems are found in about 75% of prokaryote genomes (*12*), and pAgos and BREX appear in about 10% (*7*, *13*). It has been suggested that many currently unknown defense systems reside in genomes and plasmids of nonmodel bacteria and archaea and await discovery (*2*, *14*).

Antiphage defense systems were found to be frequently physically clustered in bacterial and archaeal genomes such that, for example, genes encoding restriction enzymes commonly reside in the vicinity of genes encoding abortive infection systems and other phage-resistance systems (*14*, *15*). The observation that defense systems are clustered in genomic "defense islands" has led to the suggestion that genes of unknown function residing within such defense islands may also participate in antiphage defense (*15*, *16*). Indeed, recent studies that focused on individual genes enriched next to known defense genes resulted in the discovery of new systems that protect bacteria against phages (*7*, *9*, *17*).

Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel.
*These authors contributed equally to this work.
†Corresponding author. Email: rotem.sorek@weizmann.ac.il

Identification of putative defense gene families

We have set out to comprehensively identify new defense systems enriched within defense islands, in an attempt to systematically map the arsenal of defense systems that are at the disposal of bacteria and archaea in their fight against phages. As a first step in this discovery effort, we sought to identify gene families that are enriched near known defense systems in the microbial pangenome. For this, we analyzed 14,083 protein families (pfams) in >45,000 available bacterial and archaeal genomes (overall encoding >120 million genes). Each pfam represents a set of genes sharing a common protein domain (*18*). We calculated, for each pfam, the tendency of its member genes to reside in the vicinity of one or more known defense genes (Fig. 1, A and B) (see Methods). We further selected pfams that at least 65% of their member genes were found next to defense genes and that their member genes appeared in diverse defense contexts within different genomes (at least 10% variability) (Fig. 1C). These thresholds were selected because they capture the majority of pfams that comprise known defense systems—e.g., restriction enzymes and Abi genes (Fig. 1, B and C, and table S1) (see Methods). The resulting set of 277 candidate pfams was supplemented with 35 non-pfam gene families that were previously predicted to be associated with known defense systems (*15*), as well as 23 pfams that were predicted in the same study as putatively defensive but did not pass our thresholds, altogether yielding a list of 335 candidate gene families (table S2).

From defense genes to defense systems

Antiphage defense systems are usually composed of multiple genes that work in concert to achieve defense—for example, *cas1*, *cas2*, *cas3*, and the cascade genes in type I CRISPR-Cas systems (*19*), and the R, M and S genes in type I restriction-modification systems (*3*). Genes functioning within the same defense system are frequently encoded on the same operon, and the gene order within the operon is highly conserved among distantly related organisms sharing the same system (*3*, *7*, *9*, *16*, *19*, *20*). To investigate whether the defense-associated pfams belong to multigene systems, we used each such pfam as an anchor around which we searched for commonly associated genes (Fig. 1A). For this, we collected all the neighboring genes (10 genes from each side) from all the genomes in which members of the anchor pfam occurred and clustered these genes based on sequence homology (see Methods). We then searched for cassettes of gene clusters that, together with the anchor gene, show conserved order across multiple different genomes, marking such cassettes as candidate multigene systems (see Methods) (Fig. 1A).

The gene annotations in the resulting candidate systems were manually inspected to filter out likely false predictions. We found that 39% of the cases (129 of 335) represented nondefense, mobile genetic elements, such as transposons and integrases, that are known to colocalize with defense islands (*15*) (table S2). An additional 30% (102 of 335) represented known defense systems whose pfams were not included in our original set of known defense pfams, and 17% belonged to operons probably performing metabolic or other functions not associated with defense (fig. S1A). The remaining systems possibly represent putative new defense systems. To expand our predictions with new pfams that may be specifically enriched next to the putative new defense systems, a second prediction cycle was performed, this time adding the members of the predicted new systems to the positive defense pfam set (Fig. 1A and fig. S1B) (see Methods). Altogether, 41 candidate single-gene or multigene systems were retrieved from the two prediction cycles of this analysis (table S3). We further filtered from this set systems that were largely confined to a specific taxonomic clade (e.g., systems appearing only in cyanobacteria), resulting in a set of 28 candidate systems that showed broad phylogenetic distribution.

Experimental verification strategy

We selected two bacteria, *Escherichia coli* str. MG1655 and *Bacillus subtilis* str. BEST7003, as model organisms to experimentally examine whether the predicted systems confer defense against phages (Fig. 2A). None of the candidate new systems are naturally present in the genomes of these two bacterial strains. For each

candidate system we selected source organisms from which the system was taken and heterologously cloned into one of the model organisms. To increase the probability that the cloned system would be compatible and functionally expressed within the receiving bacterium, we selected systems from mesophilic organisms as close phylogenetically as possible to *E. coli* or to *B. subtilis* and included the upstream and downstream intergenic regions so that promoters, terminators, or other regulatory sequences would be preserved. Where possible, we took at least two instances of each system (from two different source genomes), to account for the possibility that some systems may not be active in their source organism (*21*, *22*). The DNA of each system, spanning the predicted genes and the intergenic spaces, was synthesized or amplified from the source genome and cloned into the phylogenetically closest model organism—either to *E. coli* (on a plasmid) or to *B. subtilis* (genomically integrated). As a control, we repeated the procedure with five known defense systems (instances of types I, II, and III R-M systems, a type III toxin/antitoxin system, and an abortive infection gene of the AbiH family) for which source organisms were similarly selected and cloning was performed into *B. subtilis*, as well as a sixth control composed of the recently discovered DISARM defense system (*9*) (table S4).

Altogether, we attempted to heterologously clone 61 representative instances of the 28 candidate new systems, and successful cloning was verified by whole-genome sequencing (table S4). For 27 of these 28 systems, there was at least one candidate locus for which cloning was successful, and RNA sequencing (RNA-seq) of the transformants showed that, for 26 of the systems, at least one of the candidate loci was expressed in the receiving *E. coli* or *B. subtilis* strain.

The engineered bacteria were then challenged by an array of phages consisting of 10 *B. subtilis* and six *E. coli* phages, spanning the three major families of tailed double-stranded DNA (dsDNA) phages (myo-, sipho-, and podophages), as well as one single-stranded DNA (ssDNA) phage infecting *E. coli* (Fig. 2, B and C). Measuring phage efficiency of plating (EOP) on system-containing bacteria versus control cells, we found that 9 of the 26 tested systems (35%) showed protection from infection by at least one phage (Fig. 2, B and C and figs. S2 and S3). In comparison, three of the six positive control systems showed defense, with the remaining three showing no protection against the 10 *B. subtilis* phages tested (see Discussion).

We named the nine verified new systems after protective deities from various world mythologies. These defense systems comprise between 1 and 5 genes and span between 2 and 12 kb of genomic DNA (Table 1 and table S5). Where possible, we verified system consistency by testing for phage resistance in systems where individual genes were deleted (Figs. 3 to 5 and figs. S4 and S5). We found between several hundred and several thousand representations of each of the defense systems in sequenced microbial genomes, usually with broad phylogenetic distribution (fig. S6 and tables S6 to S15). Most systems were detected in >10 taxonomic phyla, and 7 of them appear in archaea (fig. S6). Some of the systems seem to target a specific family of phages (e.g., the Thoeris system appears to specifically protect from myophages), whereas others, such as the Hachiman system, provide broader defense (Fig. 2B). The genes comprising the new systems encode many protein domains that are commonly present in antiviral systems such as CRISPR-Cas and RNA interference (RNAi), including helicases, nucleases, and nucleic acid binding domains, in addition to many domains of unknown function and also atypical domains as described below. Three of the systems contain membrane-associated proteins, as predicted by the presence of multiple transmembrane helices. Below, we focus on further functional analyses for a selected set of systems.

The Zorya defense system

The Zorya system (named after a deity from Slavic mythology) was identified based on the

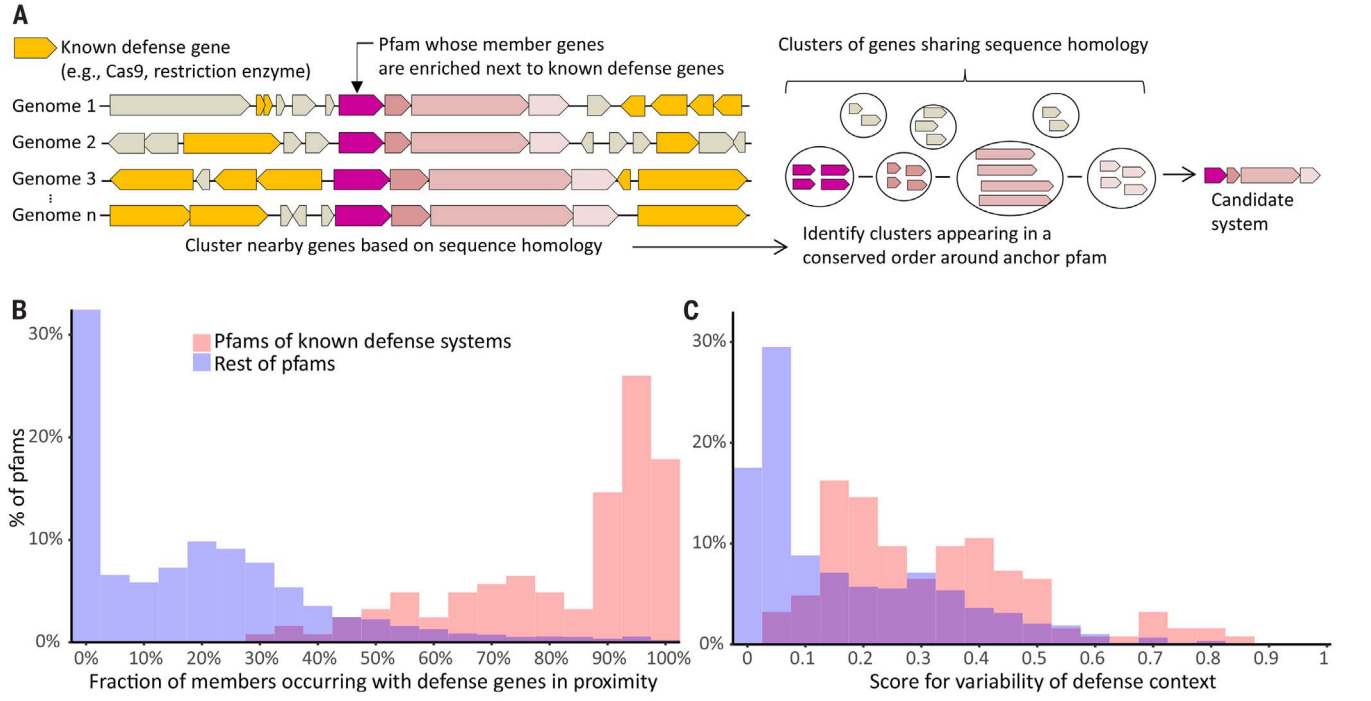


Fig. 1. Discovery of new antiphage defense systems in defense islands. (**A**) Illustration of the computational analysis employed for each pfam found to be enriched in defense islands. Pfams that are enriched in the vicinity of known defense genes are identified, and their neighboring genes are clustered based on sequence homology to identify conserved cassettes that represent putative defense systems. (**B**) Tendency of protein families to occur next to defense genes. The genomic neighborhood for each member gene in each pfam is examined, and the fraction of member genes occurring in the vicinity (10 genes on each side) of one or more known defense genes is recorded. Pink, a set of 123 pfams known to participate in antiphage defense ("positive set"); blue, the remaining 13,960 pfams analyzed in this study. (**C**) Neighborhood variability score for the analyzed pfams. Score represents the fraction of pfam members occurring in different defense neighborhoods out of total occurrences of pfam members (see Methods). Pink, the 123 positive pfams; blue, a set of 576 pfams that passed the 65% threshold for fraction of members occurring with defense genes in proximity.

enrichment of the anchor pfam15611, representing a domain of unknown function, within defense islands. Pfam15611-containing gene clusters were previously reported as genomically associated with tellurium- and stress-resistance genes (*23*). The reconstructed system is composed of the four genes *zorABCD*, overall encompassing ~9 kb of DNA, with pfam15611 being the third gene in the system (*zorC*) (Fig. 3C and Table 1). A representative Zorya operon from *E. coli* E24377A was cloned into *E. coli* MG1655 and provided 10- to 10,000-fold protection against infection by T7, SECphi27, and lambda-vir phages (Figs. 2C and 3, A and C, and fig. S3). Further searches based on homologies to the first two genes of the system, *zorA* and *zorB*, revealed a second type of Zorya, comprised of the three genes *zorABE*. A type II Zorya was cloned from *E. coli* ATCC8739 into *E. coli* MG1655 and provided defense against T7 and the ssDNA phage SECphi17 (Figs. 2C and 3, B and C, and fig. S3).

The first two genes of the Zorya system, *zorA* and *zorB*, contain protein domains sharing distant, but clear, homology with domains in *motA* and *motB*, respectively (Fig. 3C). MotA and MotB are inner membrane proteins that are part of the flagellar motor of bacteria. They assemble into a MotAB complex, which forms the stator of the flagellar motor (the static part within which the flagellar rotor swivels) (*24*). The MotAB complex also forms the proton channel that provides the energy for flagellar rotation, coupling transport of protons into the cell with the rotation (Fig. 3D) (*25*, *26*). Whereas *zorB* shares the same size and domain organization with *motB* (including the pfam13677 and pfam00691

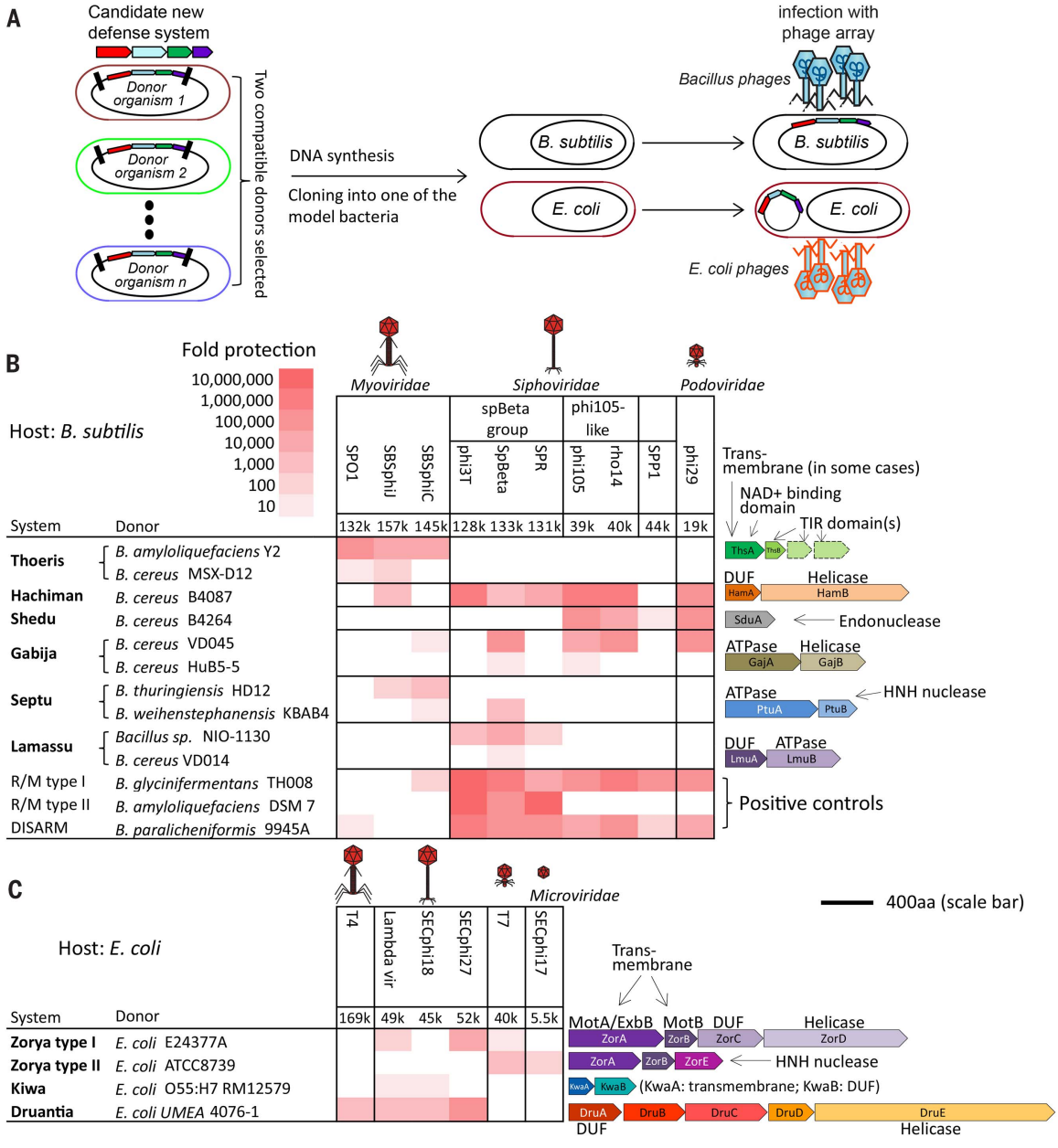


Fig. 2. Experimentally verified defense systems. (**A**) Flowchart of the experimental verification strategy. (**B**) Active defense systems cloned into *B. subtilis*. (**C**) Active defense systems cloned into *E. coli*. For (B) and (C), fold protection was measured using serial dilution plaque assays, comparing the system-containing strain to a control strain that lacks the system and has an empty vector instead. Data represent average of three replicates; see figs. S2 and S3. Numbers below phage names represent phage genome size. On the right, gene organization of the defense systems, with identified domains indicated (DUF, domain of unknown function). Gene sizes are drawn to scale; scale bar, 400 amino acids.

domains), *zorA* contains, in addition to the MotA domain (pfam01618), a long C-terminal helical domain that is sometimes identified as a methyl-accepting chemotaxis domain (COG0840). In addition to these two genes, type I Zorya contains *zorC*, a gene of unknown function, and *zorD*, which encodes a large protein (1200 amino acids) with a helicase domain that in some cases also encodes a C-terminal Mrr-like nuclease domain. Type II Zorya lacks *zorC* and *zorD* and instead contains *zorE*, a smaller gene encoding an HNH-endonuclease domain.

The gene composition of the Zorya system may point to several hypotheses as to its mechanism of action. It is possible that the system has adopted the MotAB proton channel to achieve depolarization of membrane potential upon phage infection. ZorC, ZorD, and ZorE may possibly be involved in the sensing and inactivation of phage DNA, and if phage inactivation fails, the ZorAB proton channel opens up, leading to membrane depolarization and cell death. Under this hypothesis, Zorya may be a conditional abortive infection system. Indeed, although Zorya-containing cells that were infected by phage T7 did not yield phage progeny in >80% of infection events, infection of Zorya-containing cells in liquid cultures has led to an eventual culture collapse, suggesting that Zorya-mediated defense involves death or metabolic arrest of the infected cells (fig. S7).

We further experimented with mutated forms of type I Zorya. All four genes in the system appear to be essential for its functionality, because deletion of each of the genes resulted in loss of protection from phage infection (Fig. 3E). Moreover, the activity of the ZorAB putative proton channel is necessary for the system's functionality, because point mutations in residues predicted to be critical for proton translocation through the channel (either ZorA:T147A/S184A or ZorB:D26N) yielded a nonfunctional system (Fig. 3E). Similarly, point mutations inactivating the Walker B motif of the ZorD helicase domain, predicted to prevent adenosine triphosphate (ATP) hydrolysis, resulted in loss of protection from phage infection.

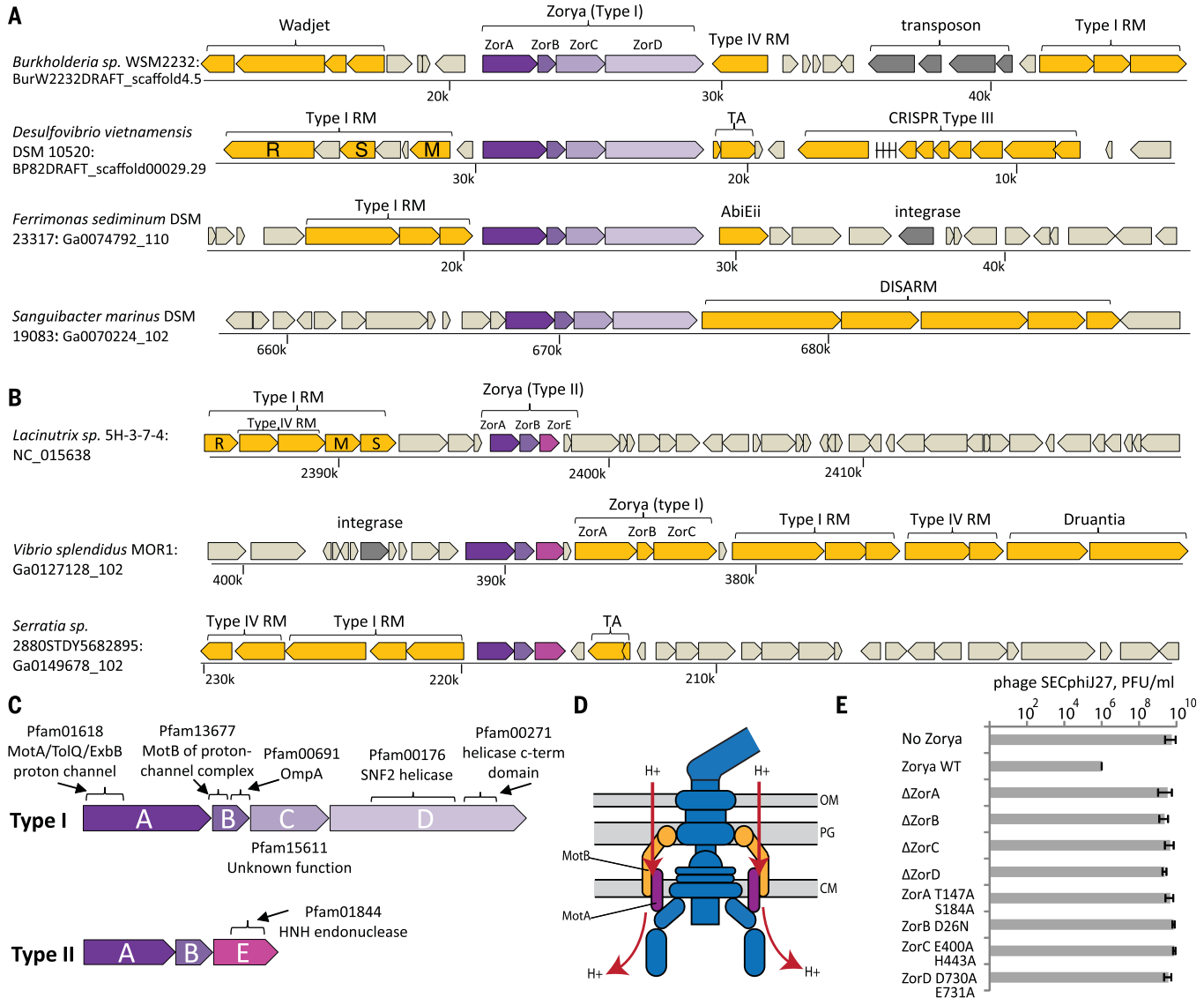


Fig. 3. The Zorya system. (**A**) Representative instances of the type I Zorya system and their defense island context. Genes known to be involved in defense are orange. Mobilome genes are in dark gray. RM, restriction-modification; TA, toxin-antitoxin; Abi, abortive infection; Wadjet and Druantia are systems identified as defensive in this study (see below). (**B**) Representative instances of the type II Zorya system. (**C**) Domain organization of the two types of Zorya. (**D**) Model of the flagellum base. The position of the MotAB complex is indicated. (**E**) EOP of phage SECphi27 infecting wild-type (WT) type I Zorya, deletion strains, and strains with point mutations. Data represent plaque-forming units per ml; average of three replicates. Error bars, mean ± SD. ZorA: T147A/S184A and ZorB:D26N are predicted to abolish proton flux; ZorC: E400A/H443A are mutations in two conserved residues in pfam15611 (EH domain) whose function is unknown (*23*); ZorD:D730A/E731A are mutations in the Walker B motif, predicted to abolish ATP hydrolysis.

We identified 1829 instances of the Zorya system within 1663 sequenced bacterial genomes, belonging to 12 phyla, marking this system as prevalent in at least 3% of sequenced bacteria (fig. S6 and table S12). We did not find the system in archaea. The system is enriched in Proteobacteria and is markedly underrepresented in Gram-positive bacteria (Firmicutes and Actinobacteria) (fig. S6), suggesting that its functionality may depend on the double-membrane organization of Gram-negative bacteria or on differences between flagellar organization of Gram-negative and Gram-positive bacteria. Because the Zorya system protects against phages that do not use flagella as their receptor [e.g., T7 (27)], Zorya protection is unlikely to stem from a receptor-masking effect.

The Thoeris defense system

Thoeris (Egyptian protective deity of childbirth and fertility) is a system that was detected based on the enrichment of pfam08937 [Toll-interleukin receptor (TIR) domain] next to known defense genes (Fig. 4A). This domain was previously reported as associated with prokaryotic argonaute genes (28). The first gene in the Thoeris system, denoted *thsA*, has an nicotinamide adenine dinucleotide (NAD)–binding domain that is sometimes annotated as sirtuin (SIR2)–like domain or Macro domain. The second gene, *thsB*, contains the TIR domain and can appear in one or more copies (Fig. 4A). In some Thoeris versions, ThsA has a multitransmembrane N-terminal domain. Two instances of this system, one from *Bacillus amyloliquefaciens* Y2 (where ThsA is predicted to be membrane-associated) and the other from *Bacillus cereus* MSX-D12 (ThsA predicted as cytoplasmic), were engineered into *B. subtilis* BEST7003 and were found to confer defense against myophages (Fig. 2B, Fig. 4, B and C, and fig. S2). Because the three myophages that we tested are very different from each other and share few homologous genes, it is possible that the Thoeris system senses or targets a general feature in the biology of myophages rather than a specific sequence or genome modification. Both Thoeris genes, *thsA* and *thsB*, are essential in the system because deletion of either of them rendered the system inactive (Fig. 4C).

Interestingly, the TIR domain is an important component of the innate immune systems of mammals, plants, and invertebrates, where it mainly serves as a connector domain that transfers the immune signal once a molecular pattern of an offensive pathogen is sensed (29). In animals, this domain frequently forms the intracellular portion of membrane-bound Toll-like receptors, whereas in plants it is often connected to intracellular R genes (30) and can also be involved in direct recognition of pathogens (31, 32). Our finding marks a common involvement of TIR domains in innate immunity across the three domains of life and implies that the ancestry of this important component of eukaryotic innate

Table 1. Composition of defense systems reported in this study.

System	Operon	Associated domains*	Domain annotations	No. of instances detected within microbes	No. (%) of genomes in which system is found	Comments
Thoeris	ThsAB	pfam13289, pfam14519, pfam08937, pfam13676	SIR2, Macro domain, TIR domain	2099	2070 (4.0%)	Membrane associated (sometimes)
Hachiman	HamAB	pfam08878, COG1204, pfam00270, pfam00271	Helicase	1781	1742 (3.4%)	
Shedu	SduA	pfam14082	Nuclease	1246	1191 (2.3%)	
Gabija	GajAB	pfam13175, COG3593, pfam00580, pfam13361 COG0210, pfam13245	ATPase, nuclease, helicase	4598	4360 (8.5%)	
Septu	PtuAB	pfam13304, COG3950, pfam13395, pfam01844	ATPase, HNH nuclease	2506	2117 (4.1%)	
Lamassu	LmuAB	pfam14130, pfam02463	SMC ATPase N-terminal domain	697	682 (1.3%)	
Zorya (type I)	ZorABCD	pfam01618, pfam13677, pfam00691, COG1360, pfam15611, pfam00176, pfam00271, COG0553, pfam04471	MotA/ExbB, MotB, helicase, Mrr-like nuclease	1173	1055 (2.1%)	Membrane associated
Zorya (type II)	ZorABE	pfam01618, pfam13677, pfam00691, COG1360, COG3183, pfam01844	MotA/ExbB, MotB, HNH nuclease	656	655 (1.3%)	Membrane associated
Kiwa	KwaAB	pfam16162	No annotated domain	934	924 (1.8%)	Membrane associated
Druantia	DruABCDE (type I) DruMFGE (type II) DruHE (III)	pfam14236, pfam00270, pfam00271, pfam09369, COG1205, pfam00145, COG0270	Helicase, methylase	1342	1321 (2.6%)	
Wadjet	JetABCD	pfam11855, pfam09660, pfam13835, pfam09661, pfam13555, pfam13558, COG4913, COG1196, pfam11795, pfam09983, pfam11796, pfam09664, COG4924	MukBEF condensin, topoisomerase VI	3173	2880 (5.6%)	

*Pfam and COG domains were assigned according to the information in the IMG database (48) and supplemented using HHpred (52).

immune systems may have stemmed from prokaryotic defense against phages.

The Thoeris system is broadly distributed in bacteria and archaea and can be detected in at least 4% of the sequenced genomes that we analyzed (2070 genomes) (table S6 and fig. S6). The TIR domain gene, *thsB*, has a strong tendency (52% of cases) to appear in multiple, diverse copies clustered around the *thsA* gene (Fig. 4A and table S6). Presence in multiple copies is typical to specificity-conferring genes in defense systems (such as the S subunit in type I R-M systems), where duplication followed by diversification serves for multiple specificities of the system (*33*–*35*). It is therefore possible that the TIR domain gene is responsible for identification of specific phage patterns, with multiple TIR domain genes serving for recognition of different phage components. Under this hypothesis, it is tempting to suggest that Thoeris is the prokaryotic ancestral form of pathogen-associated molecular pattern (PAMP) receptors.

A recent study showed that TIR domains can have enzymatic NAD^+ (oxidized form of NAD) hydrolase activities (*36*), which is in line with predictions that these domains process nucleotide derivatives (*37*). In *Caenorhabditis elegans*, this activity was shown to be involved in antifungal and antibacterial defense (*38*), whereas in animal neurons, NAD^+ hydrolysis by the SARM1 TIR domain–containing gene leads to NAD^+ depletion and generation of linear and cyclic adenosine diphosphate ribose (ADP–ribose) signaling molecules that regulate axonal degeneration (*39*). An E99A point mutation in the *B. amyloliquefaciens* Y2 ThsB protein, which aligns with the catalytic residue in the SARM1 NAD-cleaving TIR domain (fig. S8), abolished the protective activity of Thoeris (Fig. 4C). Moreover, point mutations in the ThsA NAD^+ binding site, predicted to abolish NAD^+ binding, also resulted in system inactivation (ThsA:N112A and ThsA:D100A/N115A for the *B. cereus* and *B. amyloliquefaciens* systems, respectively). These results suggest that NAD^+ binding and hydrolysis are essential for the antiphage activity of the Thoeris system.

The Druantia system

Another system worth discussing briefly is the Druantia system (named after a deity from the Gallic mythology). This system is characterized by a gene encoding a very large protein (~1800 to 2100 amino acids) containing a domain of unknown function (DUF1998) as well as a helicase signature and a Walker A/B motif suggestive of ATP utilization. This large gene is typically preceded by a set of highly variable genes with no recognizable domains or function prediction—either three genes sized 350 to 600 amino acids (type I), or two genes sized 700 to 900 amino acids (type II), or a single large gene of 1000 to 1200 amino acids (type III) (fig. S5, A and B). In some cases, type I systems are preceded by a gene annotated as DUF4338, encoding yet another domain of unknown function; and type II systems are also associated with a cytosine methylase (fig. S5, A and B). A type I system cloned from *E. coli* UMEA 4076-1 into *E. coli* MG1655 rendered the engineered strain resistant against four of the six phages tested, and by serially deleting four of the genes in this system, we verified that all four are essential for its activity (fig. S5C). Notably, DUF1998-containing genes are among the components of the recently reported DISARM (*9*) and Dpd (*40*) defense systems, where their

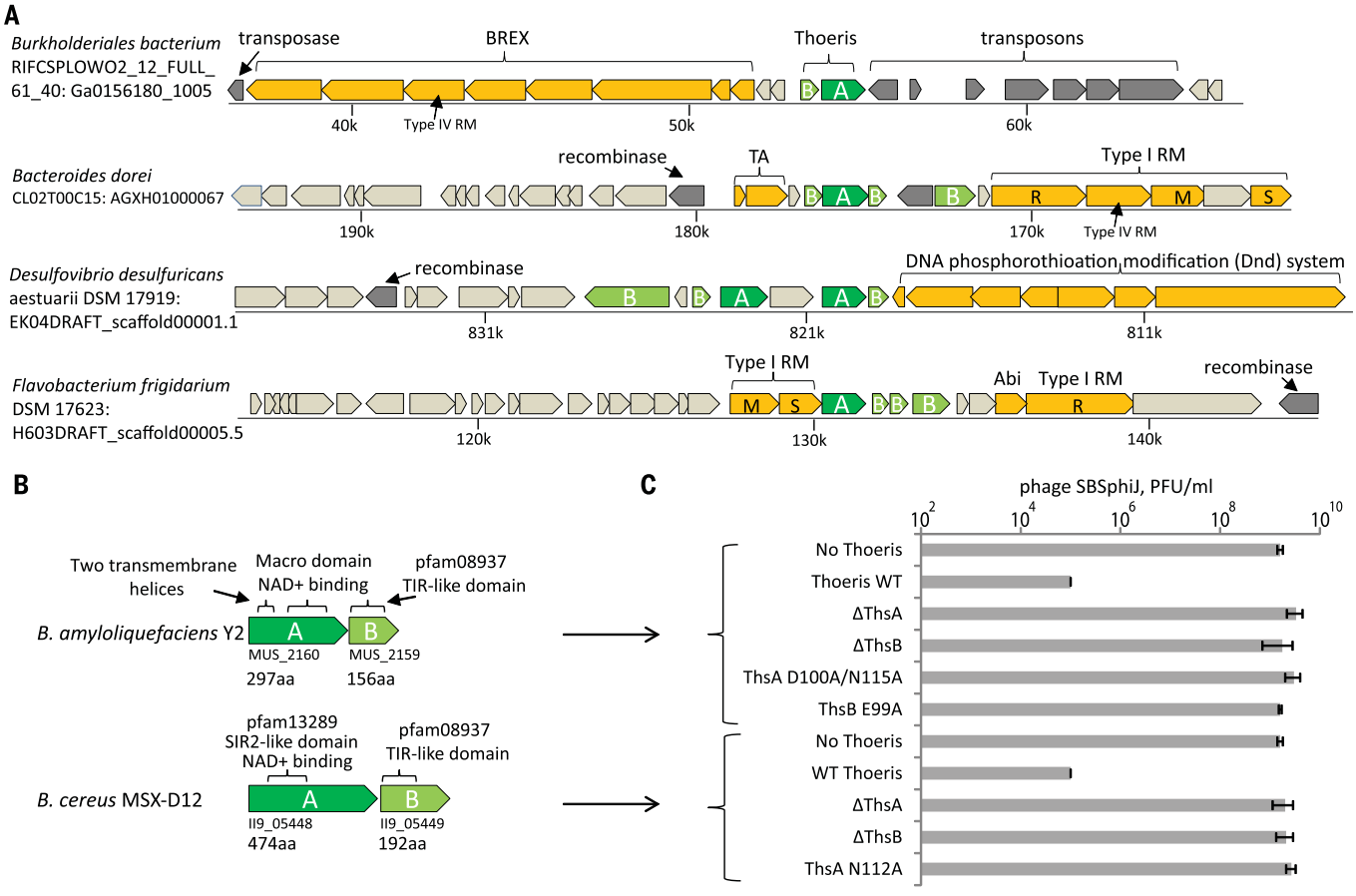


Fig. 4. The Thoeris system. (**A**) Representative instances of the Thoeris system and their defense island context. Thoeris genes *thsA* (containing NAD^+ binding domain) and *thsB* (TIR domain) are marked dark and light green, respectively. Genes known to be involved in defense are orange. Mobilome genes are in dark gray. RM, restriction-modification; TA, toxin-antitoxin; Abi, abortive infection. (**B**) The two Thoeris systems shown in this study to protect against myophages. Locus tag accessions are indicated for the individual genes. (**C**) EOP of phage SBSphiJ infection with WT and mutated versions of the *B. amyloliquefaciens* Y2 Thoeris (top set) or *B. cereus* MSX-D12 Thoeris (bottom set) cloned into *B. subtilis* BEST7003. Average of three replicates; error bars, mean ± SD.

function is also unknown. The sheer size of the Druantia system (12 kb of genomic DNA) suggests a complex function, and the near-complete absence of recognizable domains in its genes suggests a new mode of defense not shared by prokaryotic defense systems whose mechanism is currently understood.

Defense against plasmid transformation

Some of the putative defense systems that we experimentally tested did not show any antiphage activity despite being strongly associated with known defense genes. We reasoned that some of these systems may defend against other forms of foreign DNA. To test this hypothesis, we selected one such system, which we denote Wadjet (god protector of ancient Egypt), for further experimentation. Wadjet is a four-gene system, composed of the genes *jetABCD*, which is common in microbial genomes and is very frequently found next to defense genes (Fig. 5A). Three instances representing three different types of Wadjet (see below) were cloned from three separate *Bacillus* species into *B. subtilis* BEST7003. Although none of these systems provided protection against any of the 10 *Bacillus* phages in our array, all three consistently and significantly reduced transformation efficiency of the episomal plasmid pHCMC05 (Fig. 5C). These results suggest that Wadjet may be a defense system specifically targeting foreign plasmids.

We identified three different domain compositions, each encoding a different set of pfams, but all with common sequence signatures marking them as three types of Wadjet (Fig. 5B). Whereas the pfam domains of Wadjet genes are mostly defined as "domains of unknown function," structural modeling using Phyre2 (*41*) showed structural homology between JetA, JetB, and JetC and genes belonging to the housekeeping condensin system MukF, MukE, and MukB, respectively. Bacterial condensins are chromosome-organizing complexes that are responsible for DNA condensation and accurate segregation during replication (*42*), and mutations in the housekeeping condensins lead to severe defects in chromosome segregation and viability (*43*). Several versions of housekeeping condensins appear in bacterial genomes: SMC, MukBEF, and MksBEF (*44*); the Wadjet system was previously noted as a distant homolog of the MksBEF system described in *P. aeruginosa* (*45*).

Although the domain organization of the *jet*ABC genes may lead to the hypothesis that

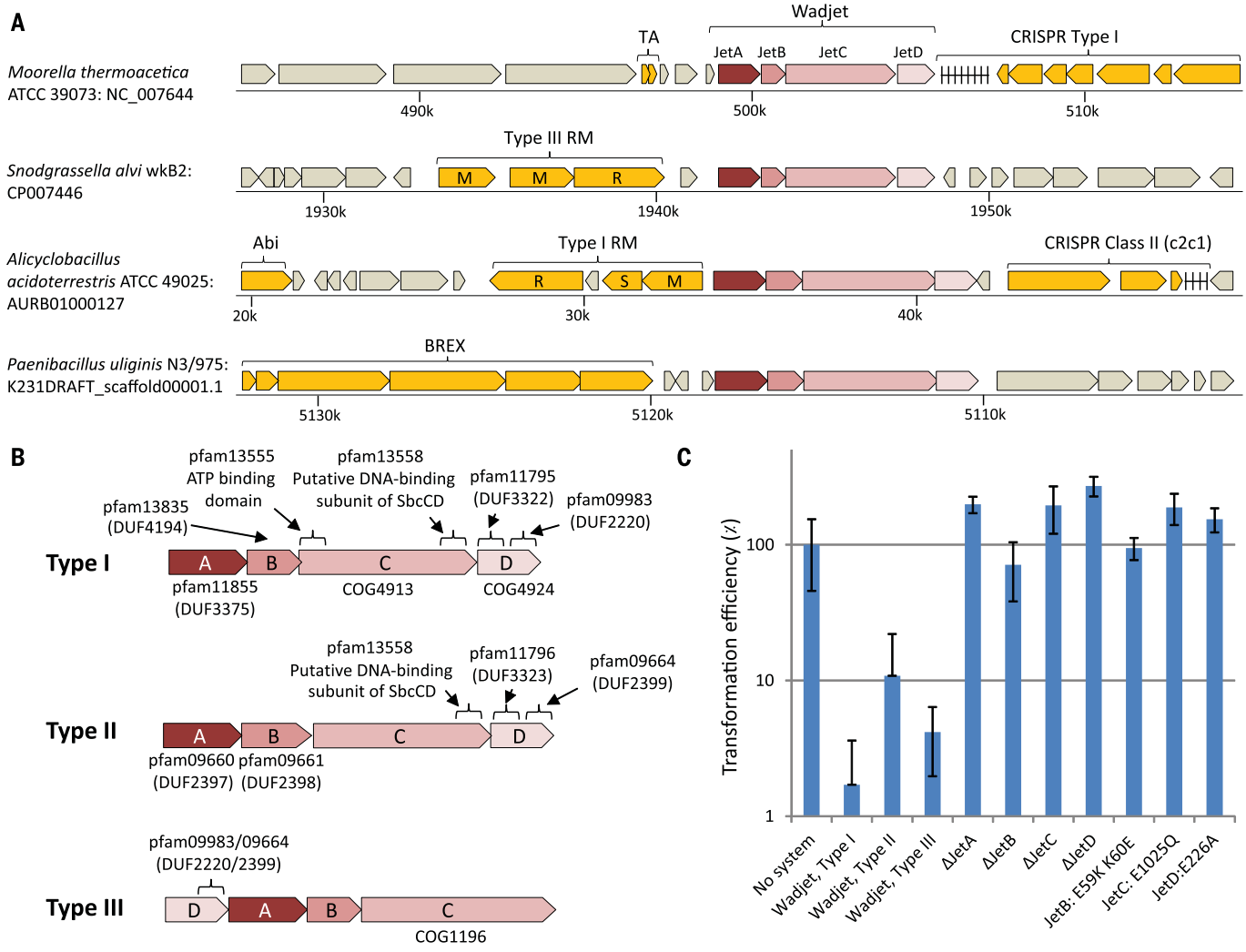


Fig. 5. The Wadjet system provides protection against plasmid transformation in *B. subtilis*. (**A**) Representative instances of the Wadjet system and their defense island context. Genes known to be involved in defense are orange. RM, restriction-modification; TA, toxin-antitoxin; Abi, abortive infection. (**B**) Domain organization of the three types of Wadjet. Pfam and COG domains were assigned according to the information in the IMG database (*48*). (**C**) Wadjet reduces plasmid transformation efficiency in *B. subtilis*. Instances of Wadjet systems were taken from *B. cereus* Q1 (type I), *B. vireti* LMG 21834 (type II), and *B. thuringiensis* serovar finitimus YBT-020 (type III) (table S4) and cloned into *B. subtilis* BEST7003. Gene deletions and point mutations are of the *B. cereus* Q1 type I Wadjet. Transformation efficiency of plasmid pHCMC05 into Wadjet-containing strains is presented as a percentage of the transformation efficiency to *B. subtilis* BEST7003 carrying an empty vector instead of the Wadjet system. Average of three replicates; error bars, mean ± SD.

Wadjet is an alternative condensin system involved in bacterial chromosome maintenance, our data imply that its role is defensive. This system is highly enriched within defense islands, undergoes extensive horizontal gene transfer, and is only sporadically found within strains of the same species, all of which is inconsistent with a core, essential role in chromosome maintenance. We hypothesize that the Wadjet system has been adapted from a MukBEF condensin ancestor to become a defense system. Possibly, the system identifies foreign plasmids and uses its condensin properties to interfere with proper plasmid segregation into daughter cells. Notably, plasmid transformation in *B. subtilis* takes place via the natural competence of this organism, during which the plasmid DNA is transformed to the cell through dedicated transporters as ssDNA (*46*). It is possible that the Wadjet system protects against rampant natural transformation or, alternatively, may specifically target ssDNA phages. However, because no ssDNA phage was reported for *B. subtilis*, we were not able to determine whether ssDNA phages are specifically blocked by the Wadjet systems cloned in *B. subtilis* BEST7003.

The Wadjet system is broadly spread in bacterial and archaeal genomes (found in ~6% of the genomes we studied), where it presents high sequence diversity (table S15 and fig. S6). Deletion of each of the four genes in type I Wadjet from *B. cereus* Q1 abolished its activity and restored plasmid transformation, indicating that each of the genes is essential for antiplasmid defense (Fig. 5C). Moreover, point mutations E59K/K60E in JetB, predicted to disrupt the MukE-MukF–like protein-protein interactions, resulted in loss of protective activity against plasmids and so has the E1025Q mutation in the Walker B motif of JetC that is predicted to abolish adenosine triphosphatase (ATPase) activity. The JetD gene, which has no homology to genes in the Muk system, has a putative topoisomerase VI domain based on structural predictions; a point mutation JetD:E226A, predicted to diminish binding of the topoisomerase VI domain to DNA, also abolished the protective activity of the system.

Discussion

Our study considerably expands the known arsenal of defense systems used by prokaryotes for protection against phages. However, our results do not yet expose the complete set of prokaryotic defense systems. Out of the 26 candidate systems we tested, nine were verified as antiphage defense systems, and an additional one showed protection against plasmids. The remaining 16, although not verified by our experiments, do not necessarily represent false predictions, as exemplified by the fact that only 50% of our positive control systems showed defense in our assays. Lack of activity of positive control systems or candidate systems could possibly stem from incompatibility of some tested systems with the recipient organism (*E. coli* or *B. subtilis*) or could be due to pseudogenization of some systems in their genome of origin. Some systems may be highly specific against a certain type of phages or foreign genetic element not represented in our phage set, whereas others may work in a specific condition not tested in our study. Clade-specific potential systems, such as those found only in archaea or cyanobacteria (table S3), were not tested in this study and can represent a more specialized defense arsenal specific only to a subset of organisms. Finally, we may have missed some true systems by falsely tagging them as belonging to the "mobilome" (table S2), as mobile genetic elements have an intimate evolutionary relationship with defense systems (*47*).

In the past, the discovery and mechanistic understanding of antiviral defense systems led to the development of important biotechnological tools. For example, the discovery of restriction enzymes resulted in a revolution in genetic engineering, and CRISPR-Cas now revolutionizes the genome editing field. Eukaryotic immune systems, such as RNAi and antibodies, have also become widely used tools. The tendency of defense systems to turn into revolutionary molecular tools stems from their intrinsic high degree of flexible molecular specificity (to differentiate between self and nonself), as well as their inherent capability to target the identified molecule. One may envision that some of the new systems we discovered, once their mechanism is deciphered, may also be adapted into useful molecular tools in the future.

Materials and methods

Computational prediction of defense systems

A set of gene families known to participate in defense

A set of pfams and COGs that are known to participate in antiphage defense was compiled based on the gene families present in table S10 from Makarova *et al.* 2011 (*15*) with the addition of pfams/COGs present in the BREX (*7*) and DISARM (*9*) antiphage systems. This set is found in table S1.

Identification of pfams enriched near defense genes

The genome sequences, gene annotations, and taxonomy annotations of all publicly available sequenced bacterial and archaeal genomes were downloaded from the NCBI FTP site (ftp.ncbi.nih.gov/genomes/genbank/bacteria/ and ftp.ncbi.nih.gov/genomes/genbank/archaea/, respectively) on April 2016. Pfam annotations for bacterial and archaeal genes were obtained from the Integrated Microbial Genomes (IMG) database (*48*) on December 2015, and cross-referenced to the genes in the genomes downloaded from NCBI using the locus_tag GenBank field. All pfams annotated in at least 20 genes ("members") across the analyzed genomes (14,083 pfams) were scanned. For each pfam, the number of member genes for which a gene having an annotation of a known defense gene family (table S1) was present in proximity (up to 10 genes upstream and 10 genes downstream) was recorded. The fraction of defense-associated members out of total members ("defense score") was calculated per pfam. A second score ("defense context variability score") was calculated for each pfam as follows: for each member gene occurring with at least one defense gene in proximity, a list of the proximal defense genes was recorded, and the fraction of unique lists out of total number of lists for that pfam represents the score (for example: if pfamX is found within 20 genes in our set, with 15 of them having Cas9 nearby and 5 having type I R-M nearby, the number of unique lists is two, and the "defense context variability score" is 2/20 = 0.1). Pfams with defense score ≥ 65% and defense context variability score ≥0.1 were taken for further analysis. This list was supplemented with 35 non-pfam gene families that were predicted to be associated with defense by Makarova *et al.* 2011 (*15*), as well as 23 pfams that were predicted in the same study but did not pass the thresholds above (table S2).

From genes to systems

Each of the putative defense-related gene families was used as an anchor to search for multigene systems, as follows. The protein coding sequences for neighboring genes (±10 genes) for all family members were clustered based on sequence homology (for example, if pfamY is found within 50 genomes in our set, the 20 neighboring genes in each genome, plus the pfamY gene in each genome, were taken—altogether 50*21 = 1050 genes to be clustered). Clustering was done with OrthoMCL software v2.0.9 (*49*) with blastp parameters [-F 'm S' -v 100,000 -b 100,000 -e 1e-5 -m 8] and with mcl v12.068 downloaded from micans.org/mcl/ (*50*, *51*) with inflation value of 1.1. When the number of blastp hits for a given anchor pfam was too large and prohibitive for OrthoMCL to generate clusters (>75 million blastp hits), a subset of genomes, containing only bacterial and archaeal genomes annotated as "complete" (rather than "draft") was used for clustering.

To detect the most prevalent genes around the anchor pfam, only the 10% largest clusters ("frequent clusters") were considered. For the sake of cluster size calculation, genes originating from the same species (derived from the strain name in the NCBI annotation) were counted as one gene, to prevent organisms for which many strains have been sequenced from inflating the cluster size. An edge between cluster(i) and cluster(j) was defined if a gene from cluster(j) followed a gene from cluster(i) in a given genome with no other genes belonging to frequent clusters found in between, with edge weight ("thickness") defined as the number of such adjacency cases. Again, edge weights were adjusted such that multiple appearances of a cluster pair originating from the same species were recorded as a single appearance. Only the 10% thickest edges were retained for further analysis. In each genome, the maximal "path" that included the anchor pfam gene and was composed of the retained (largest) clusters and the retained (thickest) edges was recorded. Such a "path," representing a set of genes appearing in a conserved order in multiple

genomes, was considered a candidate multigene system. Infrequent variations on the gene order and composition of common systems were merged into the common system if they shared at least 50% of their clusters and had less than 25% appearances than the common system. Only systems with five or more appearances from different species were further analyzed.

The domains within the gene members of each system were analyzed bioinformatically using the tools HHpred (*52*, *53*), Phyre2 (*41*), PSI-BLAST (*54*) and NCBI's Conserved Domain Database (CDD) (*55*). The systems were then manually filtered, based on this analysis, to remove (i) known defense systems whose domains did not appear in our initial set of gene families known to participate in defense; (ii) systems likely representing mobile genetic elements ("mobilome") and; (iii) systems likely participating in nondefensive functions or house-keeping systems (table S2).

A second cycle of prediction was then performed, expanding the set of "positive" gene families from table S1 to include the gene families participating in the candidate new defense systems, as well as the gene families participating in known defense systems that were previously missing from our set and detected in the first round. All pfams were again scanned and the same thresholds were applied (defense score 65%, context variability score 0.1). New pfams retrieved from the second cycle were analyzed as above to generate and annotate multigene systems.

Candidate new systems were further prioritized to select instances for experimental validations. Systems tagged as "questionable," due to uncertainty whether they represent defense genes or mobile genetic elements, were filtered out (table S3). Systems existing in only a narrow range of organisms, as well as systems that were not found in organisms phylogenetically close either to *E. coli* or *B. subtilis*, were not tested experimentally (table S3).

For system selection for experimental testing, we first attempted to select candidate systems from organisms close to *B. subtilis* as the receiving model organism, as in this organism genomic integration of large fragments of DNA is straightforward and results in a single-copy addition of the system. In case no source organisms sufficiently close to *B. subtilis* were found, we switched to *E. coli* as the model organism for experimentation.

Phylogenetic distribution analysis of new systems

For each validated defense system, several loci, including the locus that was experimentally verified, were taken as seeds for psi-Blast. psi-Blast version 2.5.0 of BLAST$^+$ (*54*, *56*), with parameters [-num_iterations 10 -max_hsps 1 -max_target_seqs 100,000 -evalue 1e-10], was performed for each protein of each system, against all microbial genomes downloaded from NCBI on April 2016. When the hits of all proteins of a system were found closely localized on a genome, spanning no more than 150% of the length of the original system, this genome was recorded as containing the system. For the Druantia and Wadjet systems, -evalue 1e-5 was used to enable detection of distant homologs. For systems with four or five genes (Zorya type I, Druantia types I and II, and Wadjet), systems were reported if at least three of their genes were identified. For the Druantia system, systems with hits to the DruE protein were retained if the DruE size was >1300 amino acids. For the Thoeris system, multiple *thsB* genes near the *thsA* gene were recorded if they were within 10 kb of genomic DNA around the identified *thsA*. Phylum for each genome was obtained using the JGI taxonomy server (https://taxonomy.jgi-psf.org/).

Experimental validation of defense systems

Cloning of candidate systems into E. coli MG1655 and B. subtilis BEST7003

A cloning shuttle vector for large fragments was constructed as previously described (*9*). The vector contains a p15a origin of replication and ampicillin resistance for plasmid propagation in *E. coli*, and *amyE* integration cassette with spectinomycin resistance for genomic integration into *B. subtilis*. The backbone of this vector was amplified using primers OGO309+OGO310, adding to it a BamHI restriction site and a terminator site upstream to the insert cloning site. The multiple cloning site of plasmid pBS1C (*57*), received from the Bacillus Genetic Stock Center (BGSC) (accession ECE257), was amplified using primers OGO311+OGO312. Both fragments were digested using AscI and BamHI, ligated using T4 ligase and transformed into *E. coli*, resulting in plasmid pSG1-rfp.

The loci of most systems were commercially synthesized and cloned, by Genscript Corp., directly into pSG1-rfp between the AscI and NotI sites of the multiple cloning site (table S4, "Cloning method" column). In one case (the type I Wadjet system) the DNA was synthesized by Gen9 (Boston, MA) with synonymous modifications to optimize GC content. In case the donor strains were readily available, the system was not synthesized but instead was directly amplified from the genomic DNA of the donor strain using KAPA HiFi HotStart ReadyMix (Kapa Biosystems KK2601) with primers as detailed in table S16. For long systems (>10,000 bases) when the donor strain was not available, the system was commercially synthesized in overlapping fragments (table S4, "Cloning method" column). Systems amplified from genomic DNA or ordered as overlapping fragments were cloned into pSG1-rfp between the AscI and NotI sites using NEBuilder HiFI DNA Assembly cloning kit (NEB E5520S). The full list of sources used for cloning the systems into our model organisms is found in table S4, including the accessions of all strains ordered.

Transformation to *B. subtilis* was performed using MC medium as previously described (*58*). MC medium was composed of 80 mM K_2HPO_4, 30 mM KH_2PO_4, 2% glucose, 30 mM trisodium citrate, 22 μg/ml ferric ammonium citrate, 0.1% casein hydrolysate (CAA), 0.2% potassium glutamate. From an overnight starter of bacteria, 10 μl were diluted in 1 ml of MC medium supplemented with 10 μl 1M $MgSO_4$. After 3 hours of incubation (37°C, 200 rpm), 300 μl was transferred to a new 15 ml tube and ~200 ng of plasmid DNA was added. The tube was incubated for another 3 hours (37°C, 200 rpm), and the entire reaction was plated on LB agar plates supplemented with 100 μg/ml spectinomycin and incubated overnight at 30°C.

For systems tested in *E. coli*, the cloned vector was transformed into *E. coli* MG1655 cells (ATCC 47076), and the resulting transformants were verified by PCR. For systems to be tested in *B. subtilis*, the cloned vector was transformed into *B. subtilis* BEST7003 cells, kindly provided previously by M. Itaya. The system was integrated into the *amyE* locus, and resulting transformants were screened on starch plates for amylase-deficient phenotype. Whole-genome sequencing was then applied to all transformed *B. subtilis* and *E. coli* clones as described in (*9*) to verify system's integrity and lack of mutations.

As a negative control for transformation into *B. subtilis*, a transformant with an empty plasmid, containing only the spectinomycin-resistance gene in the *amyE* locus, was used. As a negative control for transformation into *E. coli*, the wild-type *E. coli* MG1655 carrying an empty plasmid was used.

For strains with gene deletions and point mutations, plasmids containing systems with these deletions/mutations were commercially synthesized by Genscript. The mutated systems were transformed into *B. subtilis* and *E. coli* as described above, and clones used were fully sequenced to verify proper integration and sequence of the mutated systems.

Phage strains, cultivation, and plaque assay

The following *B. subtilis* phages were obtained from the BGSC: SPO1 (BGSCID 1P4), phi3T (BGSCID 1L1), SPβ (BGSCID 1L5), SPR (BGSCID 1L56), phi105 (BGSCID 1L11), rho14 (BGSCID 1L15), and SPP1 (BGSCID 1P7). Phage phi29 was obtained from the DSMZ (DSM 5546). Phages SBSphiJ and SBSphiC were isolated by us from mixed soil and leaves samples on *B. subtilis* BEST7003. For this, soil and leaves samples were added to a log phase *B. subtilis* BEST7003 culture and incubated overnight to enrich for *B. subtilis* phages. The enriched sample was centrifuged and filtered through 0.2 μm filters, and the filtered supernatant was used to perform double layer plaque assays as described in Kropinski *et al.* (*59*). Single plaques that appeared after overnight incubation were picked, re-isolated three times, and amplified as described below.

E. coli phages (T4, T7, and lambda-vir) were kindly provided by U. Qimron. Phages SECphi17, SECphi18, and SECphi27 were isolated as described in Wommack *et al.* (*60*) from sewage samples on *E. coli* MG1655. 0.2 μm filtered concentrated sewage samples were used to perform double layer plaque assays, individual plaques were picked, re-isolated three times, and amplified as described below.

All phages isolated by us were Illumina sequenced following a library prep using the Nextera protocol (*61*) and assembled using SPAdes v. 3.10.1 using the –careful and –cov-cutoff auto modifiers (*62*). Assembled genomes and raw reads were deposited in the European Nucleotide Archive (ENA) under study accession PRJEB23070. Phage classification was done according to sequence homology to the closest known similar phage. Phage SECphi17 (ENA ERS1981053) has a 5,538 bp genome and its closest relative is Coliphage WA3 (GenBank DQ079897.1, 66% coverage, 81% identity), indicating that it is an ssDNA phage of the *Microviridae* family. Phage SECphi18 (ENA ERS1981054) has a 44,798 bp genome and its closest relative is *Escherichia* phage Gluttony (GenBank KX534336.1, 92% coverage, 93% identity), indicating that it is a member of the *Siphoviridae* family. Phage SECphi27 (ENA ERS1981055) has a 51,811 bp genome, and its closest relative is *Escherichia* phage vB_Eco_swan01 (GenBank LT841304.1, 91% coverage, 98% identity), indicating that it is a member of the *Siphoviridae* family. Phage SBSphiJ (ENA ERS1981056) has a 156,875 bp genome, and its closest relative is *Bacillus* phage Grass (GenBank KF669652.1, 91% coverage, 95% identity), indicating that it is a member of the family *Myoviridae*. Phage SBSphiC (ENA ERS1981057) has a 144,651 bp genome, and its closest relative is Bacillus phage SP10 (GenBank AB605730.1, 94% coverage, 90% identity), indicating that it is a member of the *Myoviridae* family. *Siphoviridae* and *Myoviridae* phage morphologies were verified by electron microscopy (EM). For the EM experiments, phage lysates were blotted onto copper grids, stained using uranyl acetate 2%, and visualized in FEI Tecnai T12 transmitting electron microscope.

Phages were propagated on either *E. coli* MG1655 or *B. subtilis* BEST7003 using the plate lysate method as previously described (*63*). Lysate titer was determined using the small drop plaque assay method as previously described (*64*). Bacteria were mixed with MMB agar (LB + 0.1 mM $MnCl_2$ + 5 mM $MgCl_2$ + 5 mM $CaCl_2$ + 0.5% agar), and serial dilutions of phage lysate in MMB were dropped on top of them. After the drops dried up, plates were incubated at room temperature overnight. EOP was measured by performing small drop plaque assay with the same phage lysate on control bacteria and bacteria containing the candidate defense system, and comparing the ratio of plaque formation.

To determine the number of infective centers during infection with T7 phage of control bacteria and bacteria containing type I or type II Zorya, we used a modified version of the technique described in (*65*). Zorya-lacking *E. coli* MG1655 or Zorya-containing cells were infected with T7 phage at MOI 0.05 and incubated for 10 min at 37°C to allow adsorption. Cells with adsorbed phages were then centrifuged (1 min, 14,000 rpm) at 4°C, washed once with ice-cold MMB medium, and resuspended in 200 µl ice-cold MMB medium. Then, 100 µl aliquots of 10-fold dilutions of resuspended phage-infected cells were mixed with 100 µl of a Zorya-lacking *E. coli* MG1655 culture grown to O.D. 0.3. The mixture was plated using the double agar overlay method and infection centers (plaques) were counted after overnight incubation in room temperature.

For the liquid culture infection with T7 phage, overnight cultures of Zorya-lacking *E. coli* MG1655 or Zorya-containing cells were diluted 1:100 in MMB medium. 180 µl volumes of the diluted culture were dispersed into wells in a 96-well plate and grown at 37°C with vigorous shaking until early log phase ($O.D._{600}$ 0.3). 20 µl of T7 phage lysate were added at multiplicities of infection 0.05, 0.5 and 5 in three replicates. Optical density measurements at a wavelength of 600 nm were taken every 15 min using a TECAN Infinite 200 plate reader in a 96-well plate as previously described (*9*).

Transformation efficiency assay

Transformation was performed using the MC medium as described above. To test plasmid transformation efficiency, the episomal *Bacillus* plasmid pHCMC05 was used (*66*). Transformation efficiency was calculated by dividing the number of transformants that grew on LB plates containing 5 µg/ml chloramphenicol by the live count on LB plates.

DNA-seq and RNA-seq

DNA was extracted from bacteria using Qiagen DNeasy blood and tissue kit (Qiagen 69504). DNA libraries were constructed using the Nextera library preparation protocol as previously published (*61*). RNA-seq was performed with the NEBNext Ultra Directional RNA Library Prep Kit (NEB, E7420) according to the manufacturer's instructions with modifications as previously described (*67*). Prior to library preparation, equal amounts of extracted RNA from 3 to 7 strain samples were pooled together and processed as a single library. All libraries were sequenced using the Illumina NextSeq500. The sequencing reads were aligned to the reference genomes of *B. subtilis* BEST7003 (GenBank: AP012496) and *E. coli* MG1655 (GenBank: NC_000913), and to the plasmid sequence of each system, using Novoalign 3.02.02 (Novocraft Technologies Sdn Bhd, www.novocraft.com) with the default parameters and [-r Random]. The coverage along the reference genomes was calculated, to determine whether each system exists in the genome (DNA-seq) or expressed (RNA-seq). The pooled RNA library was sequenced to a depth of 5 million reads per sample and later aligned to the reference genomes as described.

REFERENCES AND NOTES

1. S. J. Labrie, J. E. Samson, S. Moineau, Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**, 317–327 (2010). doi: 10.1038/nrmicro2315; pmid: 20348932
2. A. Stern, R. Sorek, The phage-host arms race: Shaping the evolution of microbes. *BioEssays* **33**, 43–51 (2011). doi: 10.1002/bies.201000071; pmid: 20979102
3. R. L. Dy, C. Richter, G. P. C. Salmond, P. C. Fineran, Remarkable mechanisms in microbes to resist phage infections. *Annu. Rev. Virol.* **1**, 307–331 (2014). doi: 10.1146/annurev-virology-031413-085500; pmid: 26958724
4. M. R. Tock, D. T. Dryden, The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.* **8**, 466–472 (2005). doi: 10.1016/j.mib.2005.06.003; pmid: 15979932
5. R. Barrangou *et al.*, CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007). doi: 10.1126/science.1138140; pmid: 17379808
6. I. J. Molineux, Host-parasite interactions: Recent developments in the genetics of abortive phage infections. *New Biol.* **3**, 230–236 (1991). pmid: 1831658
7. T. Goldfarb *et al.*, BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.* **34**, 169–183 (2015). doi: 10.15252/embj.201489455; pmid: 25452498
8. D. C. Swarts *et al.*, DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**, 258–261 (2014). doi: 10.1038/nature12971; pmid: 24531762
9. G. Ofir *et al.*, DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nat. Microbiol.* **3**, 90–98 (2018). doi: 10.1038/s41564-017-0051-0; pmid: 29085076
10. J. S. Godde, A. Bickerton, The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. *J. Mol. Evol.* **62**, 718–729 (2006). doi: 10.1007/s00239-005-0223-z; pmid: 16612537
11. V. Kunin, R. Sorek, P. Hugenholtz, Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.* **8**, R61 (2007). doi: 10.1186/gb-2007-8-4-r61; pmid: 17442114
12. P. H. Oliveira, M. Touchon, E. P. C. Rocha, The interplay of restriction-modification systems with mobile genetic elements and their prokaryotic hosts. *Nucleic Acids Res.* **42**, 10618–10631 (2014). doi: 10.1093/nar/gku734; pmid: 25120263
13. D. C. Swarts *et al.*, The evolutionary journey of Argonaute proteins. *Nat. Struct. Mol. Biol.* **21**, 743–753 (2014). doi: 10.1038/nsmb.2879; pmid: 25192263
14. K. S. Makarova, Y. I. Wolf, E. V. Koonin, Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res.* **41**, 4360–4377 (2013). doi: 10.1093/nar/gkt157; pmid: 23470997
15. K. S. Makarova, Y. I. Wolf, S. Snir, E. V. Koonin, Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *J. Bacteriol.* **193**, 6039–6056 (2011). doi: 10.1128/JB.05535-11; pmid: 21908672
16. E. V. Koonin, K. S. Makarova, Y. I. Wolf, Evolutionary genomics of defense systems in archaea and bacteria. *Annu. Rev. Microbiol.* **71**, 233–261 (2017). doi: 10.1146/annurev-micro-090816-093830; pmid: 28657885
17. F. Depardieu *et al.*, A eukaryotic-like serine/threonine kinase protects *Staphylococci* against phages. *Cell Host Microbe* **20**, 471–481 (2016). doi: 10.1016/j.chom.2016.08.010; pmid: 27667697
18. R. D. Finn *et al.*, The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016). doi: 10.1093/nar/gkv1344; pmid: 26673716
19. K. S. Makarova *et al.*, An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **13**, 722–736 (2015). doi: 10.1038/nrmicro3569; pmid: 26411297
20. Y. Yamaguchi, J.-H. Park, M. Inouye, Toxin-antitoxin systems in bacteria and archaea. *Annu. Rev. Genet.* **45**, 61–79 (2011). doi: 10.1146/annurev-genet-110410-132412; pmid: 22060041
21. A. Stern, L. Keren, O. Wurtzel, G. Amitai, R. Sorek, Self-targeting by CRISPR: Gene regulation or autoimmunity? *Trends Genet.* **26**, 335–340 (2010). doi: 10.1016/j.tig.2010.05.008; pmid: 20598393
22. P. A. Hoskisson, M. C. Smith, Hypervariation and phase variation in the bacteriophage 'resistome'. *Curr. Opin. Microbiol.* **10**, 396–400 (2007). doi: 10.1016/j.mib.2007.04.003; pmid: 17719266
23. V. Anantharaman, L. M. Iyer, L. Aravind, Ter-dependent stress response systems: Novel pathways related to metal sensing, production of a nucleoside-like metabolite, and DNA-processing. *Mol. Biosyst.* **8**, 3142–3165 (2012). doi: 10.1039/c2mb25239b; pmid: 23044854
24. A. E. Baker, G. A. O'Toole, Bacteria, rev your engines: Stator dynamics regulate flagellar motility. *J. Bacteriol.* **199**, e00088–e17 (2017). doi: 10.1128/JB.00088-17; pmid: 28320878
25. D. F. Blair, H. C. Berg, The MotA protein of *E. coli* is a proton-conducting component of the flagellar motor. *Cell* **60**, 439–449 (1990). doi: 10.1016/0092-8674(90)90595-6; pmid: 2154333
26. E. R. Hosking, C. Vogt, E. P. Bakker, M. D. Manson, The *Escherichia coli* MotAB proton channel unplugged. *J. Mol. Biol.* **364**, 921–937 (2006). doi: 10.1016/j.jmb.2006.09.035; pmid: 17052729

27. V. A. González-García *et al*., Conformational changes leading to T7 DNA delivery upon interaction with the bacterial receptor. *J. Biol. Chem.* **290**, 10038–10044 (2015). doi: 10.1074/jbc.M114.614222; pmid: 25697363
28. K. S. Makarova, Y. I. Wolf, J. van der Oost, E. V. Koonin, Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol. Direct* **4**, 29 (2009). doi: 10.1186/1745-6150-4-29; pmid: 19706170
29. S. Nimma, T. Ve, S. J. Williams, B. Kobe, Towards the structure of the TIR-domain signalosome. *Curr. Opin. Struct. Biol.* **43**, 122–130 (2017). doi: 10.1016/j.sbi.2016.12.014; pmid: 28092811
30. H. Cui, K. Tsuda, J. E. Parker, Effector-triggered immunity: From pathogen perception to robust defense. *Annu. Rev. Plant Biol.* **66**, 487–511 (2015). doi: 10.1146/annurev-arplant-050213-040012; pmid: 25494461
31. T. M. Burch-Smith *et al*., A novel role for the TIR domain in association with pathogen-derived elicitors. *PLOS Biol.* **5**, e68 (2007). doi: 10.1371/journal.pbio.0050068; pmid: 17298188
32. J. L. Caplan, P. Mamillapalli, T. M. Burch-Smith, K. Czymmek, S. P. Dinesh-Kumar, Chloroplastic protein NRIP1 mediates innate immune receptor recognition of a viral effector. *Cell* **132**, 449–462 (2008). doi: 10.1016/j.cell.2007.12.031; pmid: 18267075
33. N. E. Murray, Type I restriction systems: Sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol. Mol. Biol. Rev.* **64**, 412–434 (2000). doi: 10.1128/MMBR.64.2.412-434.2000; pmid: 10839821
34. Z. Pancer, M. D. Cooper, The evolution of adaptive immunity. *Annu. Rev. Immunol.* **24**, 497–518 (2006). doi: 10.1146/annurev.immunol.24.021605.090542; pmid: 16551257
35. Y. Zhang, R. Xia, H. Kuang, B. C. Meyers, The diversification of plant *NBS-LRR* defense genes directs the evolution of microRNAs that target them. *Mol. Biol. Evol.* **33**, 2692–2705 (2016). doi: 10.1093/molbev/msw154; pmid: 27512116
36. K. Essuman *et al*., The SARM1 Toll/interleukin-1 receptor domain possesses intrinsic NAD^+ cleavage activity that promotes pathological axonal degeneration. *Neuron* **93**, 1334–1343.e5 (2017). doi: 10.1016/j.neuron.2017.02.022; pmid: 28334607
37. A. M. Burroughs, D. Zhang, D. E. Schäffer, L. M. Iyer, L. Aravind, Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. *Nucleic Acids Res.* **43**, 10633–10654 (2015). doi: 10.1093/nar/gkv1267; pmid: 26590262
38. C. Couillault *et al*., TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. *Nat. Immunol.* **5**, 488–494 (2004). doi: 10.1038/ni1060; pmid: 15048112
39. R. Fliegert, A. Gasser, A. H. Guse, Regulation of calcium signalling by adenine-based second messengers. *Biochem. Soc. Trans.* **35**, 109–114 (2007). doi: 10.1042/BST0350109; pmid: 17233614
40. J. J. Thiaville *et al*., Novel genomic island modifies DNA with 7-deazaguanine derivatives. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E1452–E1459 (2016). doi: 10.1073/pnas.1518570113; pmid: 26929322
41. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015). doi: 10.1038/nprot.2015.053; pmid: 25950237
42. T. Hirano, SMC proteins and chromosome mechanics: From bacteria to humans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360**, 507–514 (2005). doi: 10.1098/rstb.2004.1606; pmid: 15897176
43. O. Danilova, R. Reyes-Lamothe, M. Pinskaya, D. Sherratt, C. Possoz, MukB colocalizes with the oriC region and is required for organization of the two *Escherichia coli* chromosome arms into separate cell halves. *Mol. Microbiol.* **65**, 1485–1492 (2007). doi: 10.1111/j.1365-2958.2007.05881.x; pmid: 17824928
44. C. H. Haering, S. Gruber, SnapShot: SMC protein complexes, Part I. *Cell* **164**, 326–326.e1 (2016). doi: 10.1016/j.cell.2015.12.026; pmid: 26771499
45. Z. M. Petrushenko, W. She, V. V. Rybenkov, A new family of bacterial condensins. *Mol. Microbiol.* **81**, 881–896 (2011). doi: 10.1111/j.1365-2958.2011.07763.x; pmid: 21752107
46. I. Chen, P. J. Christie, D. Dubnau, The ins and outs of DNA transfer in bacteria. *Science* **310**, 1456–1460 (2005). doi: 10.1126/science.1114021; pmid: 16322448
47. M. Krupovic, K. S. Makarova, P. Forterre, D. Prangishvili, E. V. Koonin, Casposons: A new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC Biol.* **12**, 36 (2014). doi: 10.1186/1741-7007-12-36; pmid: 24884953
48. V. M. Markowitz *et al*., IMG: The Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res.* **40** (D1), D115–D122 (2012). doi: 10.1093/nar/gkr1044; pmid: 22194640
49. L. Li, C. J. Stoeckert Jr.., D. S. Roos, OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003). doi: 10.1101/gr.1224503; pmid: 12952885
50. A. J. Enright, S. Van Dongen, C. A. Ouzounis, An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584 (2002). doi: 10.1093/nar/30.7.1575; pmid: 11917018
51. S. van Dongen, C. Abreu-Goodger, Using MCL to extract clusters from networks. *Methods Mol. Biol.* **804**, 281–295 (2012). doi: 10.1007/978-1-61779-361-5_15; pmid: 22144159
52. J. Söding, A. Biegert, A. N. Lupas, The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33** (Web Server), W244–W248 (2005). doi: 10.1093/nar/gki408; pmid: 15980461
53. V. Alva, S.-Z. Nam, J. Söding, A. N. Lupas, The MPI Bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res.* **44**, W410–W415 (2016). doi: 10.1093/nar/gkw348; pmid: 27131380
54. S. F. Altschul *et al*., Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997). doi: 10.1093/nar/25.17.3389; pmid: 9254694
55. A. Marchler-Bauer *et al*., CDD: A conserved domain database for the functional annotation of proteins. *Nucleic Acids Res.* **39** (Database), D225–D229 (2011). doi: 10.1093/nar/gkq1189; pmid: 21109532
56. C. Camacho *et al*., $BLAST^+$: Architecture and applications. *BMC Bioinformatics* **10**, 421 (2009). doi: 10.1186/1471-2105-10-421; pmid: 20003500
57. J. Radeck *et al*., The *Bacillus* BioBrick Box: Generation and evaluation of essential genetic building blocks for standardized work with *Bacillus subtilis*. *J. Biol. Eng.* **7**, 29 (2013). doi: 10.1186/1754-1611-7-29; pmid: 24295448
58. G. A. Wilson, K. F. Bott, Nutritional factors influencing the development of competence in the *Bacillus subtilis* transformation system. *J. Bacteriol.* **95**, 1439–1449 (1968). pmid: 4967198
59. A. M. Kropinski, A. Mazzocco, T. E. Waddell, E. Lingohr, R. P. Johnson, in *Bacteriophages: Methods and Protocols*, M. R. J. Clokie, A. M. Kropinski, Eds. (Humana Press, NY, 2009), pp. 69–76.
60. K. E. Wommack, K. E. Williamson, R. R. Helton, S. R. Bench, D. M. Winget, in *Bacteriophages: Methods and Protocols*, M. R. J. Clokie, A. M. Kropinski, Eds. (Humana Press, NY, 2009), pp. 3–14.
61. M. Baym *et al*., Inexpensive multiplexed library preparation for megabase-sized genomes. *PLOS ONE* **10**, e0128036 (2015). doi: 10.1371/journal.pone.0128036; pmid: 26000737
62. A. Bankevich *et al*., SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012). doi: 10.1089/cmb.2012.0021; pmid: 22506599
63. L. C. Fortier, S. Moineau, Phage production and maintenance of stocks, including expected stock lifetimes. *Methods Mol. Biol.* **501**, 203–219 (2009). doi: 10.1007/978-1-60327-164-6_19; pmid: 19066823
64. A. Mazzocco, T. E. Waddell, E. Lingohr, R. P. Johnson, Enumeration of bacteriophages using the small drop plaque assay system. *Methods Mol. Biol.* **501**, 81–85 (2009). doi: 10.1007/978-1-60327-164-6_9; pmid: 19066813
65. W. D. Sing, T. R. Klaenhammer, Characteristics of phage abortion conferred in lactococci by the conjugal plasmid pTR2030. *J. Gen. Microbiol.* **136**, 1807–1815 (1990). doi: 10.1099/00221287-136-9-1807
66. M. A. Titok *et al*., *Bacillus subtilis* soil isolates: Plasmid replicon analysis and construction of a new theta-replicating vector. *Plasmid* **49**, 53–62 (2003). doi: 10.1016/S0147-619X(02)00109-9; pmid: 12584001
67. D. Dar *et al*., Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science* **352**, aad9822 (2016). doi: 10.1126/science.aad9822; pmid: 27120414

ACKNOWLEDGMENTS

We thank H. Sberro, O. Zuqert, O. Cohen, S. Mintzer, R. Shenhav, Z. Erez, D. Dar, M. Voichek, and N. Tal for useful discussion during the course of this study. We also thank M. Voichek for assistance with RNA-seq and S. Silverman for help with phage isolation. R.S. was supported, in part, by the Israel Science Foundation (personal grant 1360/16 and I-CORE grant 1796/12), the European Research Council (ERC) (grant ERC-CoG 681203), the Abisch-Frenkel Foundation, the David and Fela Shapell Family Foundation, the Benoziyo Advancement of Science grant, the Minerva Foundation, and the Pasteur-Weizmann Council. Assembled phage genomes and raw reads were deposited in the European Nucleotide Archive (ENA) under study accession PRJEB23070. Competing interests: R.S. is a scientific cofounder and advisor of BiomX Ltd. S.D., S.M., G.A., A. Leavitt, and R.S. are inventors on U.S. provisional patent application 62/586,911. S.D., S.M., G.O., and R.S. are inventors on U.S. provisional patent applications 62/512,216 and 62/512,219.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/eaar4120/suppl/DC1
Figs. S1 to S8
Tables S1 to S16

6 November 2017; accepted 28 December 2017
Published online 25 January 2018
10.1126/science.aar4120

REPORT

OPTICS

Observation of bulk Fermi arc and polarization half charge from paired exceptional points

Hengyun Zhou,[1]*† Chao Peng,[1,2]* Yoseob Yoon,[3] Chia Wei Hsu,[4] Keith A. Nelson,[3] Liang Fu,[1] John D. Joannopoulos,[1] Marin Soljačić,[1] Bo Zhen[5]‡

The ideas of topology have found tremendous success in closed physical systems, but even richer properties exist in the more general open or dissipative framework. We theoretically propose and experimentally demonstrate a bulk Fermi arc that develops from non-Hermitian radiative losses in an open system of photonic crystal slabs. Moreover, we discover half-integer topological charges in the polarization of far-field radiation around the bulk Fermi arc. Both phenomena are shown to be direct consequences of the non-Hermitian topological properties of exceptional points, where resonances coincide in their frequencies and linewidths. Our work connects the fields of topological photonics, non-Hermitian physics, and singular optics, providing a framework to explore more complex non-Hermitian topological systems.

In recent years, topological effects have been widely explored in closed and lossless systems, where the physics is characterized by a Hermitian operator that ensures a real energy spectrum and a complete, orthogonal set of eigenfunctions. This has revealed a number of previously unknown phenomena such as topologically nontrivial band structures (*1*, *2*), and promising applications including backscattering-immune transport (*3*–*5*). However, most systems, particularly in photonics, are generically non-Hermitian because of radiation into open space or material gain or loss. Non-Hermiticity enables even richer topological properties, often with no counterpart in Hermitian frameworks (*6*–*8*). One such example is the emergence of a new class of degeneracies, commonly referred to as exceptional points (EPs), where two or more resonances of a system coalesce in both eigenvalues and eigenfunctions (*9*). So far, isolated EPs in parameter space (*10*–*12*) and continuous rings of EPs in momentum space (*13*) have been studied across different wave systems because of their intriguing properties, such as unconventional transmission or reflection (*14*) and relations to parity-time symmetry (*15*), as well as their unique applications in sensing (*16*, *17*) and single-mode lasing (*18*, *19*).

We theoretically design and experimentally realize a new configuration of isolated EP pairs in momentum space, which allows us to reveal topological signatures of EPs in the band structure and far-field polarization, and to extend topological band theory into the realm of non-Hermitian systems. Specifically, we demonstrate that a Dirac point (DP) with nontrivial Berry phase can split into a pair of EPs (*20*–*22*) when radiation loss—a form of non-Hermiticity—is added to a two-dimensional (2D)–periodic photonic crystal (PhC) structure. The EP-pair generates a distinct double–Riemann sheet topology in the complex band structure, which leads to two notable consequences: bulk Fermi arcs and polarization half topological charges. First, we demonstrate that this pair of EPs is connected by an open-ended isofrequency contour—a bulk Fermi arc—in direct contrast to the common intuition that isofrequency contours are necessarily closed loops. The bulk Fermi arc here is a special topological signature of non-Hermitian effects in paired EPs and resides in the bulk dispersion of a 2D system. This is fundamentally different from the previously known surface Fermi arcs that arise from the 2D projection of Weyl points in 3D Hermitian systems. Moreover, we find experimentally that around the Fermi arc, the far-field polarization of the system exhibits a robust half-integer winding number (*23*–*25*), analogous to the orientation reversal on a Möbius strip. We show that this is a direct consequence of the topological band-switching properties across the Fermi arc connecting the EP pair and is direct experimental proof of the $\nu = \pm 1/2$ topological index associated with an EP (*8*). With comprehensive comparisons between analytical models, numerical simulations, and experimental measurements, our results are a direct validation of non-Hermitian topological band theory and present its novel application to the field of singular optics.

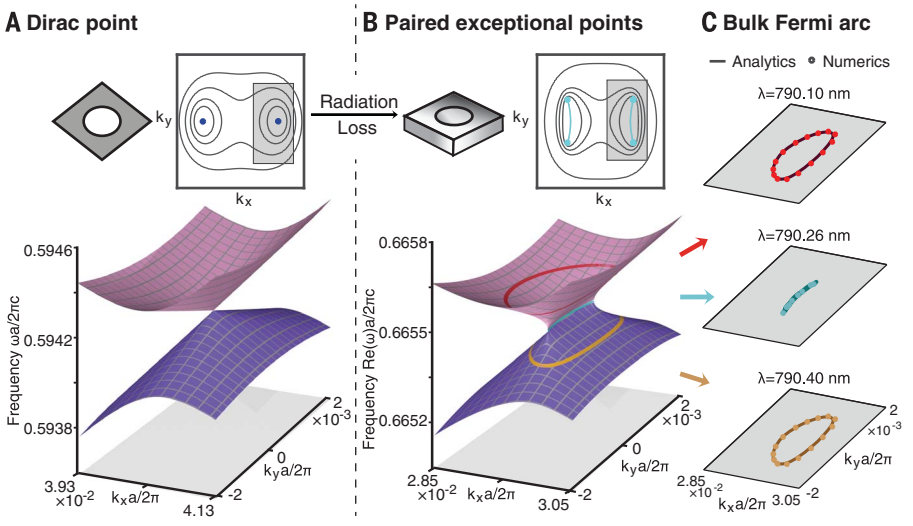


Fig. 1. Bulk Fermi arc arising from paired exceptional points split from a single Dirac point. (**A** and **B**) Illustration of photonic crystal (PhC) structures, isofrequency contours, and band structures. (A) Band structure of a 2D-periodic PhC consisting of a rhombic lattice of elliptical air holes, featuring a single Dirac point on the positive k_x axis. (B) The real part of the eigenvalues of an open system consisting of a 2D-periodic PhC slab with finite thickness, where resonances experience radiation loss. The Dirac point splits into a pair of exceptional points (EPs). The real part of the eigenvalues is degenerate along an open-ended contour—the bulk Fermi arc (blue line)—connecting the pair of EPs. (**C**) Examples of the isofrequency contours in this system, including the open bulk Fermi arc at the EP frequency (middle panel), and closed contours at higher (upper panel) or lower (lower panel) frequencies. Solid lines are from the analytical model, and circles are from numerical simulations.

[1]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]State Key Laboratory of Advanced Optical Communication Systems and Networks, Department of Electronics, Peking University, Beijing 100871, China. [3]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [4]Department of Applied Physics, Yale University, New Haven, CT 06520, USA. [5]Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA.
*These authors contributed equally to this work. †Present address: Department of Physics, Harvard University, Cambridge, MA 02138, USA.
‡Corresponding author. Email: bozhen@sas.upenn.edu

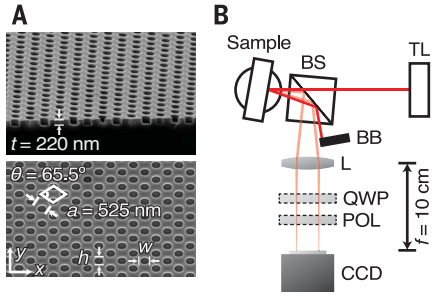


Fig. 2. Fabricated PhC slab and measurement setup. (**A**) SEM images of the PhC samples: side view (top panel) and top view (bottom panel). (**B**) Schematic of the scattering measurement setup. Vertically polarized light from a tunable continuous-wave Ti-Sapphire laser (TL) scatters off the PhC slab and is collected by a CCD camera placed at the focal plane of the lens (L). The specular reflection is blocked to ensure only scattered light is imaged. POL and QWP are used in polarimetry measurements of the scattered light. TL, tunable laser; BS, beam splitter; L, convex lens with 10-cm focal length; QWP, quarter-wave plate; POL, polarizer; BB, beam block.

Our scheme involves splitting a single DP into a pair of EPs, which directly leads to the emergence of a bulk Fermi arc. First, consider a 2D-periodic PhC with a square lattice of circular air holes introduced into a dielectric material. In this Hermitian system (no material gain, loss, or radiation loss), the crystalline symmetry (C_{4v}) ensures a quadratic band degeneracy at the center of the Brillouin zone (*26*). As this C_{4v} symmetry is broken, e.g., by shearing the structure into a rhombic lattice with elliptical holes (Fig. 1A), the quadratic degeneracy point splits into a pair of DPs situated at $(\pm k_D, 0)$ along the k_x axis. The same splitting behavior is shown in both analytical models and numerical simulations (*21*, *26*, *27*).

Next, we consider a non-Hermitian system consisting of a finite-thickness PhC slab (inset of Fig. 1B), where modes near the DP become resonances with finite lifetime because of radiative losses toward the top and bottom. Adopting the even and odd y–mirror-symmetric eigenstates at the DP as basis and taking into account the radiation losses via non-Hermitian perturbations, the effective Hamiltonian in the vicinity of the original DP at $(k_D, 0)$ can be written as (*8*, *21*)

$$H_{\text{eff}} = \omega_D - i\gamma_0 + (v_g\delta k_x - i\gamma)\sigma_z + v_g\delta k_y\sigma_x \quad (1)$$

with complex eigenvalues of

$$\omega_\pm = \omega_D - i\gamma_0 \pm \sqrt{(v_g^2\delta k^2 - \gamma^2) - 2i\gamma v_g\delta k_x} \quad (2)$$

Here, $\sigma_{x,z}$ are Pauli matrices, ω_D is the DP frequency, $(\delta k_x, \delta k_y)$ is the momentum displacement from $(k_D, 0)$, and $\delta k^2 = \delta k_x^2 + \delta k_y^2$. Meanwhile, $\gamma_0 \pm \gamma$ are the radiation decay rates of the even and odd y–mirror-symmetric modes, taking into account that the two modes have different coupling strengths to the continuum; v_g is the group velocity describing the dispersion around the DP, which for simplicity is here chosen to be the same along all directions [see (*26*) for the general case]. The real part of the complex eigenvalues $\omega_\pm$ characterizes the resonance frequency, whereas the imaginary part represents the linewidth of the resonance.

The eigenvalue spectrum exhibits a pair of EPs at $(\delta k_x, \delta k_y) = (0, \pm\gamma/v_g)$, near the original DP, when the square-root term in Eq. 2 vanishes and the two eigenvectors coalesce (Fig. 1B). The existence of such EPs is topologically robust against continuous changes in the Hamiltonian (*8*, *26*) and does not rely on any symmetries or fine-tuning. Furthermore, this pair of EPs is connected in momentum space by an open-ended arc—a bulk Fermi arc, along which the real part of the complex eigenvalues is degenerate at ω_D (Fig. 1C, middle panel). Although sharing features similar to previously studied Fermi arcs—both are open-ended isofrequency contours—our bulk Fermi arc resides at one frequency in the bulk dispersion rather than on the surface of a 3D Hermitian system and originates from non-Hermiticity rather than from the presence of Weyl points. As the frequency ω decreases from above ω_D, the closed isofrequency contour at ω shrinks (Fig. 1C, top panel), eventually turning into the open Fermi arc when $\omega = \omega_D$ (Fig. 1C, middle panel), and expands out again into a closed contour at even lower frequencies (Fig. 1C, bottom panel). Taken together, the band structure around the EPs forms a double–Riemann sheet topology (Fig. 1B). This originates from the complex square-root term in the dispersion in Eq. 2, which, depending on the sign choice of the square root, results in two sheets. The two eigenvalues continuously evolve on each sheet, and their real parts become degenerate along an open-ended curve—the bulk Fermi arc. From a different point of view, this can also be understood as a nodal arc in the bulk band structure, in analogy to nodal lines in semimetal systems (*2*). We further verify the existence of bulk Fermi arcs in realistic PhC slab structures via numerical simulations (Fig. 1C, circles), showing a good agreement with analytical results (Fig. 1C, solid lines) (*26*). The extent of the bulk Fermi arcs can be tuned by engineering the band structure and coupling rates to the continuum.

To experimentally demonstrate the bulk Fermi arc, we use interference lithography to fabricate PhC slabs in Si_3N_4 (refractive index $n = 2.02$, thickness $t = 220$ nm) on top of a silica substrate ($n = 1.46$). The PhC structure consists of rhombic unit cells with side length $a = 525$ nm, unit cell angle $\theta = 65.5°$, and elliptical air holes with long-axis length $w = 348$ nm and short-axis length $h = 257$ nm (*26*). Scanning electron microscope (SEM) images of the fabricated samples are shown in Fig. 2A. The structure is immersed in an optical liquid with refractive index matched to that of the silica substrate to create an up-down symmetric environment.

We performed angle-resolved scattering measurements (setup shown in Fig. 2B) to image isofrequency contours of the sample. The PhC sample is illuminated with a tunable continuous-wave Ti:Sapphire laser that is vertically polarized, while scattered light—arising from natural fabrication imperfections of the sample—is collected with a charge-coupled device (CCD) camera placed at the focal plane of a convex lens with 10-cm focal length. Because of resonant enhancement, the scattered light will have strongest intensity only along directions where the underlying resonances share the same frequency as the pump laser, and thus the isofrequency contours of the sample are directly imaged onto the CCD (*26*, *28*, *29*). Compared to previous measurement techniques based on the reflection spectrum (*13*), this scattering method enables fast and direct extraction of band information, without fitting to any specific models. To show the full shrinking-and-reexpanding feature of the isofrequency contours around the bulk Fermi arc, the laser wavelength is tuned from 794 nm down to 788 nm at steps of ~0.2 nm. Furthermore, the polarization at each point along a given isofrequency contour is determined through polarimetry measurements, by optionally inserting a quarter-wave plate and a polarizer in front of the CCD (*26*).

At a few representative wavelengths around the Fermi arc, the numerical results of isofrequency contours (Fig. 3A) obtained from simulating extracted structural parameters are plotted against the experimental results (Fig. 3B), showing good agreement with each other. Here, for better comparison, the numerical results are offset by 0.5 nm relative to the experiments. [See (*26*) about possible reasons for this wavelength offset and the full set of isofrequency contours measured at different wavelengths.] To focus on the bulk Fermi arc, we highlight the region of interest in both panels, where the isofrequency contours clearly demonstrate the shrinking and reexpanding behavior. As shown in Fig. 3, as the wavelength decreases from 794.0 nm, the corresponding isofrequency contour shrinks (top two rows), and eventually becomes an open-ended arc at 791.0 nm (middle row), consistent with our previous theoretical predictions in Fig. 1C. As the wavelength is further decreased down to 789.5 nm and 788.7 nm, the arc expands out into closed contours again (bottom two rows). The bending feature of the contours is a result of higher-order terms in the band dispersion (*26*). The open contour at 791.0 nm (middle row) is a clear, direct observation of the bulk Fermi arc.

So far, we have shown one direct consequence of the unique double–Riemann sheet topology near paired EPs—the bulk Fermi arc. Next, we demonstrate another consequence: half-integer topological charges in the polarization configuration, which also serve as a direct experimental proof of the $\nu = \pm 1/2$ topological index of an EP. These topological charges describe the direction (clockwise or counterclockwise) and number of times the polarization vector winds around a point or line singularity in the optical field, and in our particular system, we observe a robust

180° winding around the Fermi arc, corresponding to a half-integer topological charge.

To fully reconstruct the far-field polarization configurations of the resonances, we perform polarimetry measurements by recording the intensity of isofrequency contours after passing through six different configurations of polarizers and/or waveplates (*26*). Although the incoming light is vertically polarized, the scattered light at each point along the contour is, in general, elliptically polarized, reflecting the polarization state of its underlying resonance. Taking points X and Z in Fig. 4A as examples: After passing through a vertical polarizer, the scattered light is weak (strong) at point X(Z); whereas after a horizontal polarizer, the relative intensity of the scattered light switches between points X and Z. This clearly shows that the far field of the underlying resonance at point X(Z) is mostly horizontally (vertically) polarized.

Examples of the fully reconstructed spatial polarizations (blue ellipses) at representative points along the 794-nm isofrequency contour (red solid line) are shown in the top panel of Fig. 4B, which agree well with numerical results (Fig. 4B, bottom panel). Furthermore, both experimental and numerical results show 180° winding of the polarization long axis, as illustrated by the green arrows in Fig. 4B: As the momentum point starts from point X, traverses the full contour in the counterclockwise direction, and returns to point X, the polarization long axis flips direction by rotating 180° in the clockwise direction—corresponding to a $-\frac{1}{2}$ topological charge being enclosed in the loop. These results thus indicate that the far-field emission from our PhC is a vector-vortex beam with half-integer topological charge, in stark contrast to the integer vector beams realized in photonic crystal surface–emitting lasers (*24*).

We now explain the fundamental connections between the half-integer topological charges observed in the far-field polarization and the half-integer topological index of an EP (*8*), manifested as its mode-switching property (*26*). Along the k_x axis, the two bands forming the EP pair in our system have orthogonal linear polarizations due to the y–mirror symmetry: One is horizontal (e.g., mode X in Fig. 4C), whereas the other is vertical (e.g., modes Z and W). As we follow a closed path in momentum space X → Y → Z → W that encircles one of the EPs in the counterclockwise direction, the initial eigenstate X (horizontally polarized) on the top sheet adiabatically evolves into state Z (vertically polarized) and eventually into final state W (vertically polarized) on the bottom sheet, owing to the mode-switching topological property of the EP (*10*–*12*). The switching behavior of the eigenmodes—from X to W—directly follows from their eigenvalue swapping behavior on the complex plane (*26*). Equivalently, one complex eigenvalue winds around the other one by half a circle, thus implying that the topological index of an EP is a half-integer. The orthogonal nature between the polarizations at X and Z, arising from the mode-switching property of the EP, guarantees

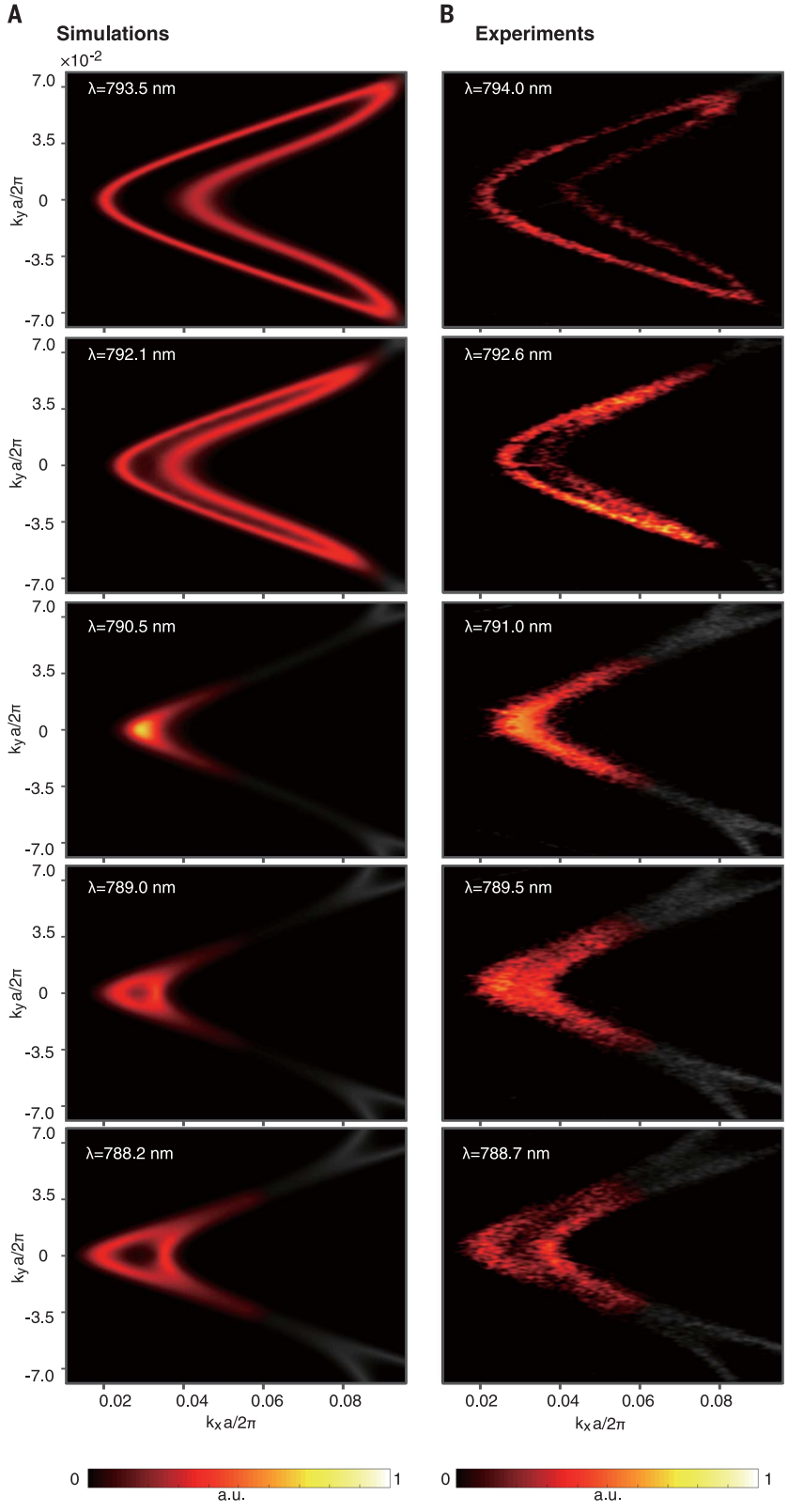


Fig. 3. Experimental demonstration of a bulk Fermi arc. (**A**) Numerically simulated spectral density of states and (**B**) experimentally measured isofrequency contours at five representative wavelengths. The bulk Fermi arc appears at 791.0 nm (middle row), when the isofrequency contour becomes open-ended. The regions of interest are highlighted in all panels to emphasize the shrinking (top two rows) and reexpanding (bottom two rows) feature of isofrequency contours near the bulk Fermi arc. The numerical results are offset by 0.5 nm for better comparison.

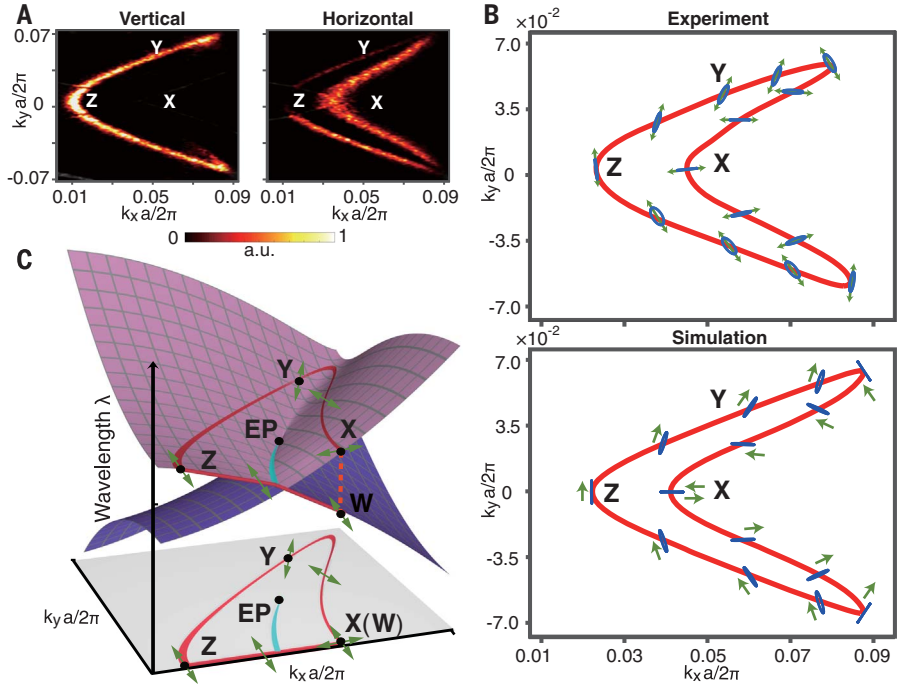


Fig. 4. Experimental demonstration of polarization topological half charges around the bulk Fermi arc. (**A**) Intensity of scattered light at 794 nm after passing through a vertical (horizontal) polarizer is plotted in the left (right) panel, showing that point X(Z) is mostly horizontally (vertically) polarized. (**B**) Experimental reconstruction (top) and numerical simulation (bottom) of the full polarization information, showing the polarization ellipses (blue ellipses) as well as their long-axis directions (green arrows) along an isofrequency contour (red line). As shown by the green arrows in the bottom panel, the polarization long axis exhibits a $-\frac{1}{2}$ topological charge. (**C**) Schematic illustration of the mode switching (X to W) in the band structure, along a loop enclosing an EP (X-Y-Z-W), as a result of the double–Riemann sheet topology. This mode-switching behavior directly leads to the half-integer topological index of an EP and the half-charge polarization winding.

a $(n+\frac{1}{2})\pi$-rotation ($n \in \mathbb{Z}$) of the polarization vector along half the contour. Again using the y–mirror symmetry, the full isofrequency contour will accumulate twice the rotation angle, to a combined $(n+\frac{1}{2}) \times 2\pi$-rotation, corresponding to a half-integer topological charge of $n+\frac{1}{2}$.

We have thus shown the intimate connection between polarization vector winding in singular optics and the double–Riemann sheet topology of paired EPs. Our experimental demonstration—generating half-integer vector-vortex beams directly from the topological properties of EPs—not only distinguishes our study from the previously known integer topological charges of polarization around bound states in the continuum (*30*), but also proves the nontrivial topology of EPs.

We have demonstrated that the topological properties of paired EPs endow the band structure and far-field emission with unique features, manifested as the emergence of bulk Fermi arcs and polarization half topological charges. Our structure also provides an easily realizable method to create half-integer vector-vortex beams (*25*) at a wide range of frequencies. Future prospects leveraging the topological landscape around paired EPs may enable PhC lasers with exotic emission profiles (*24*), such as twisted Möbius strips. The isolated EPs found in our structure also provide a straightforward platform for studying the influence of EPs and their topology on light-matter interactions, such as modified Purcell factors for spontaneous emission enhancement and nonlinear optics generation. Our observation of bulk Fermi arcs and polarization half charges extends the existing framework of topological physics in closed systems into a new regime involving open systems and provides a platform for the future exploration of non-Hermitian topological physics in general wave systems, ranging from photonic and acoustic to electronic and polaritonic systems.

REFERENCES AND NOTES

1. A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, A. K. Geim, *Rev. Mod. Phys.* **81**, 109–162 (2009).
2. N. P. Armitage, E. J. Mele, A. Vishwanath, *Rev. Mod. Phys.* **90**, 015001 (2018).
3. M. Hasan, C. Kane, *Rev. Mod. Phys.* **82**, 3045–3067 (2010).
4. X. L. Qi, S. C. Zhang, *Rev. Mod. Phys.* **83**, 1057–1110 (2011).
5. L. Lu, J. D. Joannopoulos, M. Soljačić, *Nat. Photonics* **8**, 821–829 (2014).
6. N. Moiseyev, *Non-Hermitian Quantum Mechanics* (Cambridge Univ. Press, 2011).
7. D. Leykam, K. Y. Bliokh, C. Huang, Y. D. Chong, F. Nori, *Phys. Rev. Lett.* **118**, 040401 (2017).
8. H. Shen, B. Zhen, L. Fu, arXiv:1706.07435 [cond-mat.mes-hall] (22 June 2017).
9. W. D. Heiss, *J. Phys. Math. Gen.* **45**, 444016 (2012).
10. C. Dembowski *et al.*, *Phys. Rev. Lett.* **86**, 787–790 (2001).
11. J. Doppler *et al.*, *Nature* **537**, 76–79 (2016).
12. H. Xu, D. Mason, L. Jiang, J. G. E. Harris, *Nature* **537**, 80–83 (2016).
13. B. Zhen *et al.*, *Nature* **525**, 354–358 (2015).
14. Z. Lin *et al.*, *Phys. Rev. Lett.* **106**, 213901 (2011).
15. C. M. Bender, S. Boettcher, *Phys. Rev. Lett.* **80**, 5243–5246 (1998).
16. H. Hodaei *et al.*, *Nature* **548**, 187–191 (2017).
17. W. Chen, Ş. Kaya Özdemir, G. Zhao, J. Wiersig, L. Yang, *Nature* **548**, 192–196 (2017).
18. H. Hodaei, M.-A. Miri, M. Heinrich, D. N. Christodoulides, M. Khajavikhan, *Science* **346**, 975–978 (2014).
19. L. Feng, Z. J. Wong, R.-M. Ma, Y. Wang, X. Zhang, *Science* **346**, 972–975 (2014).
20. O. N. Kirillov, A. A. Mailybaev, A. P. Seyranian, *J. Phys. Math. Gen.* **38**, 5531–5546 (2005).
21. H. Zhou, "Tailoring light with photonic crystal slabs: From directional emission to topological half charges," thesis, Massachusetts Institute of Technology (2016).
22. V. Kozii, L. Fu, arXiv:1708.05841 [cond-mat.mes-hall] (19 August 2017).
23. M. V. Berry, M. R. Dennis, *Proc. R. Soc. A Math. Phys. Eng. Sci.* **459**, 1261–1292 (2003).
24. E. Miyai *et al.*, *Nature* **441**, 946 (2006).
25. T. Bauer *et al.*, *Science* **347**, 964–966 (2015).
26. See supplementary materials.
27. Y. D. Chong, X.-G. Wen, M. Soljačić, *Phys. Rev. B* **77**, 235125 (2008).
28. E. C. Regan *et al.*, *Sci. Adv.* **2**, e1601591 (2016).
29. L. Shi, H. Yin, X. Zhu, X. Liu, J. Zi, *Appl. Phys. Lett.* **97**, 251111 (2010).
30. B. Zhen, C. W. Hsu, L. Lu, A. D. Stone, M. Soljačić, *Phys. Rev. Lett.* **113**, 257401 (2014).

ACKNOWLEDGMENTS

We thank T. Savas for fabrication of the samples. We also thank S. Johnson, E. Mele, L. Lu, Y. Shen, S. Skirlo, S. Liu, J. Y. Lee, F. Machado, N. Rivera, G. Zhang, and S. Moore for helpful discussions. Research was supported in part by the Army Research Office through the Institute for Soldier Nanotechnologies under contract no. W911NF-13-D-0001 [photon management for developing nuclear-thermophotovoltaic (TPV) and fuel-TPV mm-scale systems]. Research was also supported as part of the S3TEC, an Energy Frontier Research Center funded by the U.S. Department of Energy under grant no. DE-SC0001299 (for fundamental photon transport related to solar TPVs and solar thermoelectrics), and by the Air Force Research Lab under contract FA8650-16-D-5403. H.Z. acknowledges support from the Undergraduate Research Opportunities Program at Massachusetts Institute of Technology. B.Z. acknowledges the Air Force Office of Scientific Research Young Investigator Program under award number FA9550-18-1-0133. C.P. acknowledges support from the National Natural Science Foundation of China under grant nos. 61575002 and 61320106001, and the China Scholarship Council. C.W.H. acknowledges support from the National Science Foundation under grant NSF DMR-1307632. Y.Y. and K.A.N. were supported in part by Skoltech as part of the Skoltech Next Generation Program. H.Z. and B.Z. conceived the idea. H.Z. performed the analytical calculations and numerical simulations. H.Z., B.Z., C.P., and Y.Y. conducted the experiments and analyzed the data. H.Z. and B.Z. wrote the manuscript, with input from all authors. B.Z., M.S., and J.D.J. supervised the research. All authors contributed to the analysis and discussion of the results. All data needed to evaluate the conclusions are present in the paper and/or the supplementary materials. Additional data related to this paper may be requested from the authors.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1009/suppl/DC1
Supplementary Text
Figs. S1 to S4
References (*31*–*38*)
Movie S1

17 September 2017; accepted 19 December 2017
Published online 11 January 2018
10.1126/science.aap9859

OPTICS

Ideal Weyl points and helicoid surface states in artificial photonic crystal structures

Biao Yang,[1]* Qinghua Guo,[1,2]* Ben Tremain,[3]* Rongjuan Liu,[4]* Lauren E. Barr,[3] Qinghui Yan,[4] Wenlong Gao,[1] Hongchao Liu,[1] Yuanjiang Xiang,[2] Jing Chen,[5] Chen Fang,[4] Alastair Hibbins,[3]† Ling Lu,[4]† Shuang Zhang[1]†

Weyl points are the crossings of linearly dispersing energy bands of three-dimensional crystals, providing the opportunity to explore a variety of intriguing phenomena such as topologically protected surface states and chiral anomalies. However, the lack of an ideal Weyl system in which the Weyl points all exist at the same energy and are separated from any other bands poses a serious limitation to the further development of Weyl physics and potential applications. By experimentally characterizing a microwave photonic crystal of saddle-shaped metallic coils, we observed ideal Weyl points that are related to each other through symmetry operations. Topological surface states exhibiting helicoidal structure have also been demonstrated. Our system provides a photonic platform for exploring ideal Weyl systems and developing possible topological devices.

Topology is the mathematics of conserved properties under continuous deformations, and the recent study of band topologies is yielding a suite of fascinating interface transport phenomena that include one-way propagation of energy and previously unknown relativistic behavior. The two-dimensional honeycomb lattice is the most studied in the exploration of topological phenomena. Made famous by graphene (*1*), the energy-momentum dispersion in a honeycomb system is linear, and the crossings of bands in energy-momentum space are known as Dirac points. The transport of the quasiparticles around these points is massless, and this remarkable transport behavior is associated with the "hidden" symmetry associated with its two identical sublattices. Weyl points are the characteristic of an analogous phenomenon when the lattice is extended to three dimensions (*2*–*5*). In electronic systems, materials exhibiting Weyl points are known as Weyl semimetals, and Weyl fermion is the solution to the massless Dirac equation. Each Weyl point can be assigned an integer "charge" based on its chirality, known as the Chern number, and much like magnetic monopoles, Weyl points are only ever found in pairs of opposite charge. Just like Dirac points, these Weyl points also exist in photonic systems, but unlike Dirac points, they can only exist once either (or both) time-reversal or space-inversion symmetry of the crystal is broken. To date, Weyl points of various forms have been proposed and realized in several boson or fermion systems (*2*–*4*, *6*–*13*). Among them, the presence of surface state arcs as one of the fingerprints of Weyl systems has been observed.

However, demonstration of more fundamental topological features of Weyl points—such as the helicoidal dispersion, which yields the open Fermi arcs of topological surface states (*14*)—has been hindered by the complicated configuration of energy bands at the Weyl energy. Moreover, some realistic and innovative device applications critically depend on a simple embodiment of Weyl systems (*5*). Thus, an ideal Weyl system (*15*–*17*) has attracted much attention because in such systems, all Weyl nodes are symmetry-related, residing at the same energy with a large momentum separation and devoid of nontopological bands in a sufficiently large energy interval.

Although Weyl degeneracies can be readily found by breaking either time-reversal or

[1]School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK. [2]International Collaborative Laboratory of 2D Materials for Optoelectronic Science and Technology of Ministry of Education, Shenzhen University, Shenzhen 518060, China. [3]Electromagnetic and Acoustic Materials Group, Department of Physics and Astronomy, University of Exeter, Stocker Road, Exeter EX4 4QL, UK. [4]Institute of Physics, Chinese Academy of Sciences/Beijing National Laboratory for Condensed Matter Physics, Beijing 100190, China. [5]School of Physics, Nankai University, Tianjin 300071, China.
*These authors contributed equally to this work.
†Corresponding author. Email: a.p.hibbins@exeter.ac.uk (A.H.); linglu@iphy.ac.cn (L.L.); s.zhang@bham.ac.uk (S.Z.)

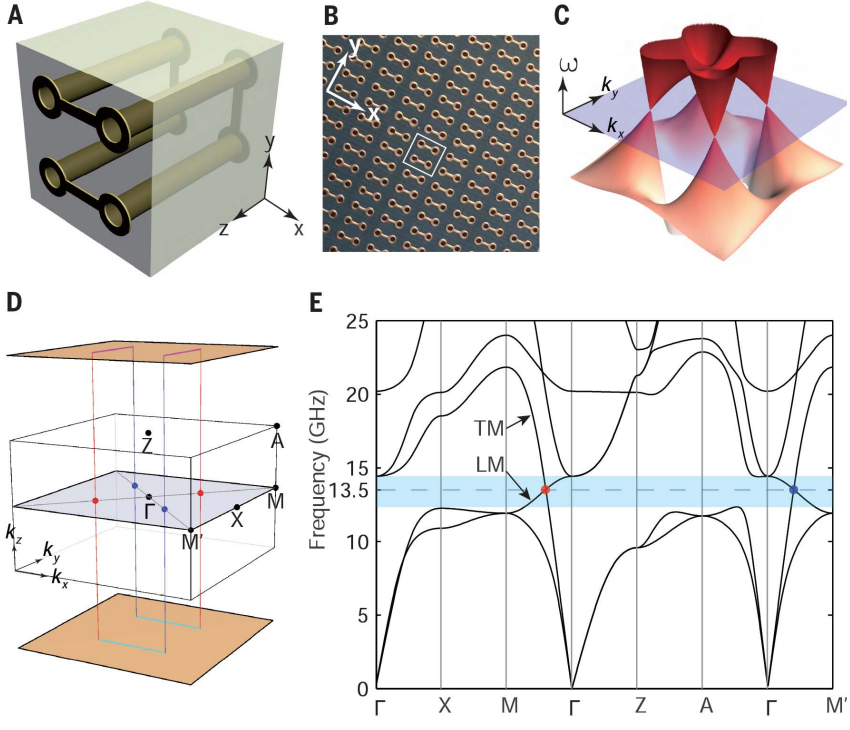


Fig. 1. Structure and band topology of the ideal photonic Weyl meta-crystal. (**A**) Schematic of a saddle-shaped metallic inclusion, which has noncentrosymmetric D_{2d} point group symmetry, embedded in a dielectric (dielectric constant of 2.2 ± 2% at 10 GHz). Here, period $a_x = a_y = a = 3$ mm and $a_z = 4.5$ mm. (**B**) Photograph of the top surface of the sample, fabricated with printed circuit board technology by etching 3-mm-thick, double-sided, copper-clad (0.035 mm-thick) dielectric laminates. A 1.5-mm-thick "blank" layer spaces each pair of printed layers so as to prevent electrical connection between the metallic coils. The bulk sample is assembled by stacking (1.5 + 3)–mm bilayers in the *z* direction. The unit cell is indicated by the white square. (**C**) Four type-I Weyl points reside on the same energy, as indicated by the blue plane with respect to $k_z = 0$. (**D**) Bulk and surface BZ with four Weyl points located along the Γ–M directions. Top (magenta) and bottom (cyan) topological surface-state arcs are shown schematically. (**E**) CST Microwave Studio (CST) simulated band structure along high-symmetry lines. The blue shaded area highlights the energy window where the ideal Weyl points (red and blue points) reside. Longitudinal mode (LM) and transverse mode (TM) are labeled.

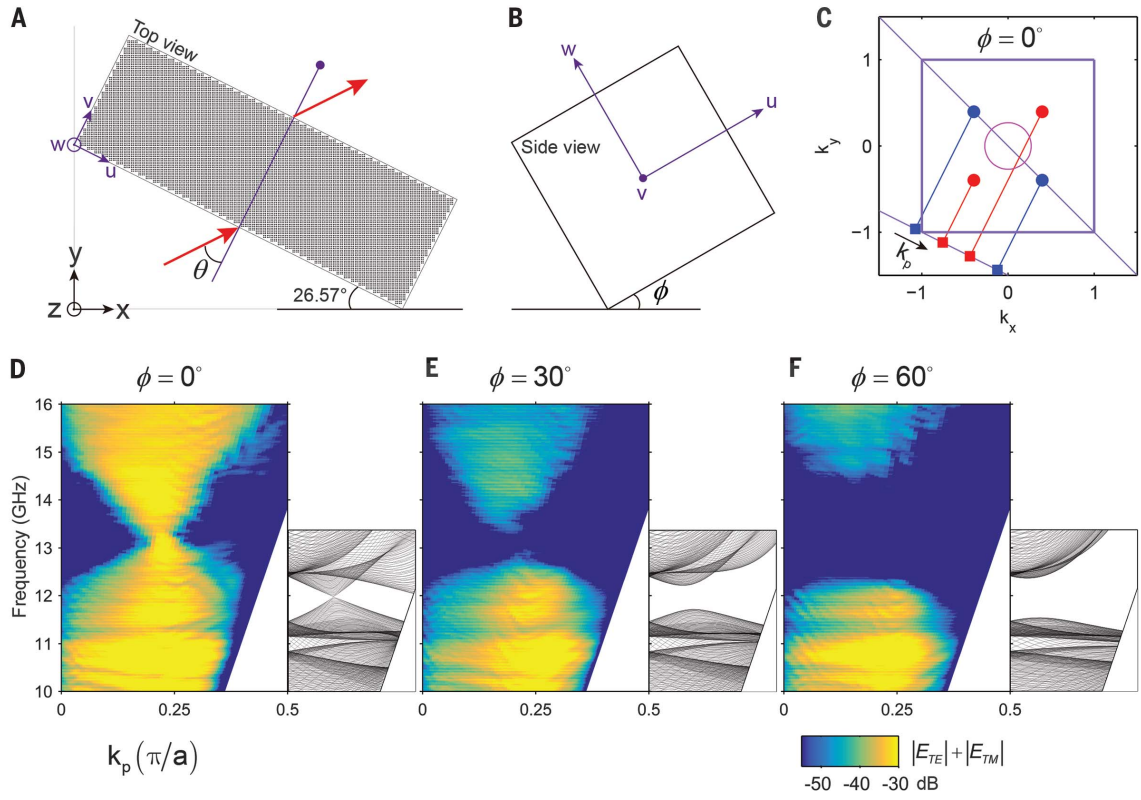


Fig. 2. Angle-resolved transmission measurement of the ideal Weyl system. (**A** and **B**) Schematic view of the sample fabricated with crystal cutting angle of 26.57°. Top and side views are illustrated. θ and ϕ are the rotation angles defined around the *z* and *v* axes, respectively. (**C**) Projection of Weyl points in momentum space with respect to the global coordinates (*x*, *y*, *z*) when ϕ = 0°. First BZ is indicated by the purple square. Magenta circle indicates the equifrequency contour of vacuum at 13.5 GHz. k_p is the in-plane component of the incident wave vector through a projection onto the sample surface (*u*-*w* surface) (*9*). (**D**, **E**, and **F**) The band projections with ϕ = 0°, 30°, and 60°, respectively. The experimental and simulated results are shown at left and right, respectively.

inversion symmetry (*5*), or both, the experimental realization of a truly ideal Weyl system has not yet been reported. Here, we explore the microwave response of a three-dimensional photonic crystal composed of metallic inclusions (termed a "meta-crystal") in order to realize an ideal Weyl system protected by D_{2d} point symmetry. Our meta-crystal exhibits four Weyl points at the same energy, the minimum number allowed in the presence of time-reversal symmetry. By placing an excitation point-source on one surface of the crystal, and scanning the near fields on the opposite surface, we observed the intriguing helicoidal structure of topological surface states: a physical representation of Riemann surface generated by a multivalued function (*14*).

Our meta-crystal design offers an ideal platform for the investigation of various unconventional physics in Weyl systems. The symmetry of the studied meta-crystal belongs to the simple tetragonal lattice with symmorphic space group $P\bar{4}m2$ (no. 115). The basis comprises a saddle-shaped connective metallic coil (Fig. 1, A and B) that possesses D_{2d} ($\bar{4}2m$ in Hermann-Mauguin notation) point group symmetry. The system has no spatial inversion. These metallic elements support localized electromagnetic resonances with current distributions that can be expanded into multipolar modes (*18*). In an effective medium model (supplementary materials section 4 and fig. S1) (*19*), these resonances collectively exhibit a bi-anisotropic effect, leading to a directionally dependent chirality response (*20*). Here, the unavoidable crossings between the longitudinal mode (LM) with negative dispersion and the transverse electric mode (TM) with positive dispersion along Γ–M result in the formation of a type-I Weyl point (Fig. 1C and fig. S2) (*19*, *21*). Analysis via the irreducible representation of the point group shows that these two modes belong to two different classes, with eigenvalues ±1 of C_2 rotation along Γ–M (Fig. 1D), where level repulsion is forbidden (supplementary materials section 5 and fig. S3) (*19*). The other three Weyl points are obtained after application of the D_{2d} symmetry operation. For instance, three twofold rotation symmetries (C_2 and $2C_2'$) combined with time-reversal symmetry guarantee that these four Weyl nodes are located on the Γ–M at the same frequency, at which application of two mirror symmetries (σ_x and σ_y) reverses the corresponding topological charges. The simulated band structure along high-symmetry lines is shown in Fig. 1E (as defined in Fig. 1D) in the Brillouin zone (BZ), where a pair of Weyl points resides at the same frequency. As such, these Weyl degeneracies appear in a relatively large energy window (~2.1 GHz around the frequency of the Weyl point) (Fig. 1E, blue shaded region) that is also devoid of other bulk bands and hence unequivocally facilitates their experimental identification.

The linear band crossings of the bulk states around the Weyl point are confirmed with angle-resolved transmission measurements (*9*). In order to couple energy across the meta-crystal, the momentum of the bulk states must be matched to the in-plane momentum of an incident wave; a sample with special crystal cutting is fabricated (Fig. 2A), in which the crystal orientation forms an angle of 26.57° with one of the cutting boundaries. Compared with the global axis, *xyz*, a local coordinate, *uvw*, is defined. The length (along *u*), width (along *v*), and height (along *w*) of the sample are 300, 100, and 300 mm, respectively. Two angles, θ and ϕ (Fig. 2, A and B), are scanned to obtain the angle-resolved transmission spectrum. With this specific crystal cutting, when ϕ = 0°, Weyl points in BZ are projected along the scan wave vector k_p, which is related to θ, as indicated in Fig. 2C. Obviously, two of the projected Weyl points are located within the light circle (magenta circle) at the Weyl frequency (13.5 GHz). Thus, even a plane wave illuminated directly from air onto the sample can address these two Weyl points. Comparisons between the simulation and experiment results are shown in Fig. 2, D, E, and F. In Fig. 2D, with ϕ = 0°, a linear gapless energy dispersion is obtained, and the density of states vanishes at the Weyl frequency because of the absence of other bulk states at the same frequency in an ideal Weyl system. After rotating the sample to ϕ = 30° and 60° around the *v* axis, a complete gap is observed as expected.

Another direct manifestation of the topological aspects of a Weyl system is the exotic topological surface states taking the form of arcs connecting the topologically distinct bulk states. Following a closed contour around an end of the arcs, one moves between the lower (valence) and upper (conduction) bands (*14*), which is a direct consequence of the chiral characteristic of Weyl

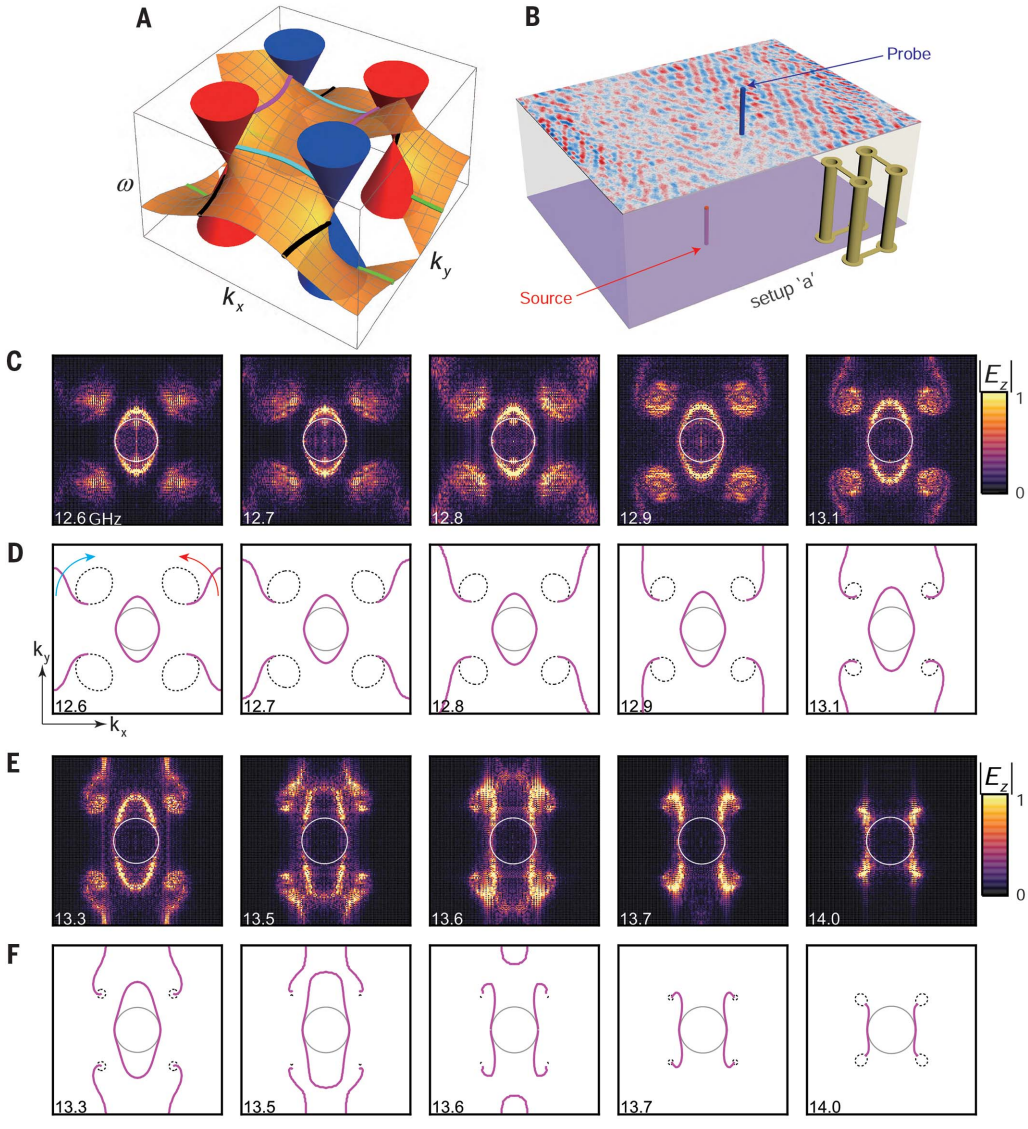


Fig. 3. Experimental observation of helicoidal structure of topological surface states. (**A**) Schematic illustration of helicoid surface states in an ideal Weyl system, with four Weyl points within the surface BZ, plotted by using the Jacobi elliptic function. The arcs of different colors represent the evolution of equifrequency arcs connecting Weyl points of opposite Chern numbers. (**B**) Transmitted near-field scanning system (setup "a"), where the source (red) is positioned on the bottom surface center. (**C** and **E**) Equifrequency contour ($|E_z|$) measured by using setup "a" from 12.6 to 14.0 GHz. (**D** and **F**) Bulk (black dashed) and surface (magenta solid) states simulated with, correspondingly. Anticlockwise (red) and clockwise (cyan) arrows indicate the surface arc rotation directions with increasing frequency corresponding to positive and negative Weyl nodes, respectively. The central solid circle indicates the air equifrequency contour. The plotted range for each panel is $[-\pi/a, \pi/a]^2$.

nodes, as schematically shown in Fig. 3A. It is well known that the gapless surface states of Weyl crystals take the form of helicoid Riemann surfaces (*14*), where the bulk Weyl points correspond to the poles and zeros adopting the sign of their respective Chern numbers. Recently, it was shown that topological surface states of double Weyl systems can be analytically expressed, across the entire BZ, as the double-periodic Weierstrass elliptic function (*22*). Because the Weierstrass elliptic function has one second-order pole and one second-order zero, it is not the most fundamental expression of the Weyl surface states. Here, we show that our ideal-Weyl meta-crystal of four Weyl points has surface states whose dispersion is topologically equivalent to the argument of Jacobi elliptic function cn(z,m) of two poles and two zeros on the complex plane. cn(z,m) is a meromorphic function with periods 4K(m) and 4K(1 − m), where K is the complete elliptic integrals of the first kind. For our system, the mapping is given by $\omega(k_x, k_y) \sim \arg\{\mathrm{cn}[(k_x - k_y)/2 + (k_x + k_y)i/2, 1/2]\}$, as plotted in Fig. 3A.

The helicoidal structure of the surface arcs was probed by using the transmitted near-field scanning configuration, with the excitation source located at the center of the bottom layer of the meta-crystal stack (Fig. 3B, setup "a"), where the detecting probe can raster-scan the top surface so as to map out both the bulk and surface modes. Another configuration (fig. S4B, setup "b"), in which the excitation source is positioned at the edge or corner of the top surface, is also used to identify the surface states. These two setups provide complementary information for the observation of helicoid surface states. In all near-field measurements, we set the scanning step as 1 mm (a/3), providing a large surface momentum space in the range of $(-3\pi/a, 3\pi/a)^2$ after the Fourier transformation. The helicoid structure of the surface arc was experimentally measured and numerically simulated and is presented as a series of equifrequency contours between 12.6 and 14.0 GHz (Fig. 3, C and E, in experiment, and Fig. 3, D and F, in simulation).

As shown in Fig. 3, C and D, at 13.1 GHz, which is below the Weyl frequency, the Fourier transformation of the experimentally measured field distribution shows the presence of four symmetrically displaced elliptical bulk states with the same size located along the diagonal directions. We clearly observed two surface arcs running across the BZ boundaries and connecting the neighboring bulk states with opposite topological charges. In the vicinity of the air equifrequency contour (circle), there exists a surface ellipse. The surface ellipse joins and reroutes the surface arc at higher frequencies (Fig. 3, E and F). Indeed, the surface ellipse and surface arcs together form the same unified helicoid surface in the dispersion of the surface states.

With increasing frequency, the top surface arc that emerged from the Weyl node with positive and negative topological charge rotates anticlockwise and clockwise, respectively. The observed rotation of the helicoid surface state around a Weyl node can therefore be used to detect the chirality of the Weyl node (*23*). As mentioned above, at

lower frequencies each surface arc connects between the bulk states through the BZ boundary, whereas the surface ellipse expands gradually with increasing frequency. Between 13.5 and 13.6 GHz, the surface arc and surface ellipse connect with each other and then transition into a new configuration: a direct surface arc connecting between the bulk states within the BZ, and a surface ellipse centered at its edge. The evolution of the surface arc configuration across the measured frequency range matches topologically with that described by the Jacobi elliptic function shown in Fig. 3A. At the frequency of 14.3 GHz, the surface arcs appear to be linear (fig. S4F) (*19*), leading to nearly diffractionless propagation of the surface wave in the real space (fig. S4C) (*19*). Slightly away from the Weyl frequency, the equi-energy contour of the bulk state consists of four very small spheres enclosing the Weyl points. It is expected that the interference between them results in a chessboard-like interference pattern in real space, which is experimentally confirmed as a spatial frequency filter (fig. S5) (*17*, *19*). In addition, the dimensional reduction from ideal Weyl points to graphene-like dispersion is shown in fig. S6 (supplementary materials section 8) (*19*). We also analyzed that the presence of dielectric loss of our system does not affect the existence of the Weyl points (supplementary materials section 9) (*19*).

The designed ideal Weyl system presented here opens up opportunities for studying intriguing physics and offers a prototype platform for realistic device applications. In photonics, besides the topologically nontrivial surface states supported by Weyl materials, the diverging Berry curvature (*5*) close to Weyl points provides a new degree of freedom in controlling the transport of optical wave packets and may lead to the observation of a gigantic Hall effect for light (*24*). Furthermore, any phenomena related to the conical dispersion of the light cone may be observed around Weyl points, such as diverging and diminishing scattering cross sections (*25*). The vanishing density of states at Weyl frequencies also provides a robust platform for controlling light-matter interaction when emitters are embedded inside photonic Weyl materials.

REFERENCES AND NOTES

1. A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, A. K. Geim, *Rev. Mod. Phys.* **81**, 109–162 (2009).
2. X. Wan, A. M. Turner, A. Vishwanath, S. Y. Savrasov, *Phys. Rev. B* **83**, 205101 (2011).
3. A. A. Burkov, L. Balents, *Phys. Rev. Lett.* **107**, 127205 (2011).
4. G. Xu, H. Weng, Z. Wang, X. Dai, Z. Fang, *Phys. Rev. Lett.* **107**, 186806 (2011).
5. N. P. Armitage, E. J. Mele, A. Vishwanath, arXiv:1705.01111 [cond-mat.str-el] (2017).
6. S.-Y. Xu *et al*., *Science* **349**, 613–617 (2015).
7. H. Weng, C. Fang, Z. Fang, B. A. Bernevig, X. Dai, *Phys. Rev. X* **5**, 011029 (2015).
8. A. A. Soluyanov *et al*., *Nature* **527**, 495–498 (2015).
9. L. Lu *et al*., *Science* **349**, 622–624 (2015).
10. L. Huang *et al*., *Nat. Mater.* **15**, 1155–1160 (2016).
11. W.-J. Chen, M. Xiao, C. T. Chan, *Nat. Commun.* **7**, 13038 (2016).
12. J. Noh *et al*., *Nat. Phys.* **13**, 611–617 (2017).
13. B. Yang *et al*., *Nat. Commun.* **8**, 97 (2017).
14. C. Fang, L. Lu, J. Liu, L. Fu, *Nat. Phys.* **12**, 936–941 (2016).
15. J. Ruan *et al*., *Nat. Commun.* **7**, 11136 (2016).
16. J. Ruan *et al*., *Phys. Rev. Lett.* **116**, 226801 (2016).
17. L. Wang, S.-K. Jian, H. Yao, *Phys. Rev. A* **93**, 061801 (2016).
18. T. Kaelberer, V. A. Fedotov, N. Papasimakis, D. P. Tsai, N. I. Zheludev, *Science* **330**, 1510–1512 (2010).
19. Materials and methods are available as supplementary materials.
20. Q. Guo, W. Gao, J. Chen, Y. Liu, S. Zhang, *Phys. Rev. Lett.* **115**, 067402 (2015).
21. M. Xiao, Q. Lin, S. Fan, *Phys. Rev. Lett.* **117**, 057401 (2016).
22. T. Zhang *et al*., arXiv:1705.07244 [cond-mat.mtrl-sci] (2017).
23. Q. Ma *et al*., *Nat. Phys.* **13**, 842–847 (2017).
24. M. Onoda, S. Murakami, N. Nagaosa, *Phys. Rev. Lett.* **93**, 083901 (2004).
25. M. Zhou *et al*., *Nat. Commun.* **8**, 1388 (2017).

ACKNOWLEDGMENTS

This work was financially supported by the European Research Council Consolidator Grant (Topological), Horizon 2020 Action Project grant 734578 (D-SPA), and Leverhulme Trust (grant RPG-2012-674). S.Z. acknowledges support from the Royal Society and Wolfson Foundation. B.Y. acknowledges support from China Scholarship Council (grant 201306110041). Q.G. acknowledges the financial support of the National Natural Science Foundation of China (grant 11604216). Y.X. acknowledges support from the National Natural Science Foundation of China (grant 61490713). L.E.B. and A.H. acknowledge financial support from the Engineering and Physical Sciences Research Council of the United Kingdom (grant EP/L015331/1). C.F. was supported by the National Key Research and Development Program of China under grant 2016YFA0302400 and by the National Natural Science Foundation of China (NSFC) under grant 11674370. L.L. was supported by the National Key R&D Program of China under grants 2017YFA0303800 and 2016YFA0302400 and by NSFC under project 11721404. Near-field scanning data were collected by a vector network analyzer controlled with *xyz*-stage at G31 at the Department of Physics and Astronomy, University of Exeter, UK. All data created during this research are openly available from the University of Exeter's institutional repository at http://hdl.handle.net/10871/30744.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1013/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S6
References (*26*, *27*)

4 October 2017; accepted 19 December 2017
Published online 11 January 2018
10.1126/science.aaq1221

ORGANIC CHEMISTRY

Selective formation of γ-lactams via C–H amidation enabled by tailored iridium catalysts

Seung Youn Hong,* Yoonsu Park,* Yeongyu Hwang, Yeong Bum Kim, Mu-Hyun Baik,† Sukbok Chang†

Intramolecular insertion of metal nitrenes into carbon-hydrogen bonds to form γ-lactam rings has traditionally been hindered by competing isocyanate formation. We report the application of theory and mechanism studies to optimize a class of pentamethylcyclopentadienyl iridium(III) catalysts for suppression of this competing pathway. Modulation of the stereoelectronic properties of the auxiliary bidentate ligands to be more electron-donating was suggested by density functional theory calculations to lower the C–H insertion barrier favoring the desired reaction. These catalysts transform a wide range of 1,4,2-dioxazol-5-ones, carbonylnitrene precursors easily accessible from carboxylic acids, into the corresponding γ-lactams via sp^3 and sp^2 C–H amidation with exceptional selectivity. The power of this method was further demonstrated by the successful late-stage functionalization of amino acid derivatives and other bioactive molecules.

The catalytic oxidation of C–H bonds is the most desirable way of converting readily available raw feedstocks to useful, value-added commodity chemicals. One such reaction highly sought after in pharmaceutical as well as materials chemistry is the direct nitrogenation of aliphatic C–H bonds (*1*–*4*). An effective general method for carrying out these C–N coupling reactions is to first convert the nucleophilic amino functionalities to more reactive, electrophilic nitrene precursors that are subsequently used as reaction partners in metal-catalyzed C–H amination reactions (Fig. 1A). The initial demonstration of this chemistry was reported by Breslow in 1983 (*5*), wherein reactive hypervalent ylides acted as sulfonylnitrene precursors in the Fe(III)- or Rh(II)-catalyzed oxathiazolidine synthesis. Major advances were achieved in the early 2000s when Du Bois and others found elegant ways of generating reactive nitrogen ylide species (*6*–*9*), which could be used to prepare a variety of amide products of high synthetic utility (*7*). In addition, organic azides have been identified as productive nitrene precursors to indoles (*10*, *11*), indolines (*12*), and pyrrolidines (*13*). Recently, hydroxylamine derivatives were elegantly used as an effective handle for synthesizing aza-arenes in N-unprotected form (*14*). Despite these advances, cyclic amides such as lactams that are valuable scaffolds in synthetic and medicinal applications could not be obtained directly through a C–H amidation strategy. Considering how successful C–N coupling techniques have been in general, the absence of such a method for preparing lactams is puzzling. In principle, carbonylnitrenes generated in situ might allow for direct construction of a cyclic amide scaffold.

As shown in Fig. 1B, catalytic reactions are believed to proceed through a key metal-nitrenoid species, which inserts into aliphatic C–H bonds to generate the corresponding azaheterocyclic products. The main reason that C–H amidation is ineffective for lactam synthesis lies in the intrinsic instability of the putative carbonylnitrene intermediate, which may easily decompose and form isocyanates via a Curtius-type rearrangement (Fig. 1C, left). This instability is well documented for acyl azides that were explored as synthetic precursors under photolytic, thermolytic (*15*), and transition-metal catalysis conditions (*16*, *17*). As a result, the general consensus is that acyl nitrenes are unfit to serve as amide sources in C–H insertion processes (*18*).

We envisioned that this paradigm might be challenged if the decomposition pathway could be blocked and the desired C–H amidation of the nitrenoid could be engineered to be faster (Fig. 1C, right). Previously, we discovered that 1,4,2-dioxazol-5-ones, which can be readily obtained from abundant carboxylic acids, are versatile substitutes for acyl azides in related C–N coupling reactions (*19*); thus, we imagined that these substrates may provide a nitrene-based route to lactam synthesis (*20*). Seeking to develop a catalyst capable of C–H activation and acylnitrene formation while suppressing the Curtius-type degradation, we were drawn to an iridium complex stabilized by an electron-donating cyclopentadienyl ligand. Cyclometallated Cp*Ir(III) complexes are widely known to facilitate C–H activation under ambient conditions. These reactions proceed via an inner-sphere imido insertion into an iridium-carbon bond, where acyl azides or analogous dioxazoles serve as the nitrogen source (*21*, *22*). The key intermediate is thought to be a high-valent Ir(V)-carbonylimido species (*22*), which undergoes a reductive C–N coupling to deliver the corresponding Ir(III)-acylamido product. We hypothesized that the half-sandwich Ir(III) complex may be able to form the acylnitrenes and promote C–H insertion via an outer-sphere mechanism (Fig. 1D), envisioning that catalytic C–N coupling in these systems need not be limited to inner-sphere mechanisms. We were encouraged by the fact that the Cp*Ir(III) platform did not produce detectable amounts of isocyanates during other C–N coupling processes; hence, we anticipated that the competitive decomposition pathway might be controlled by proper tuning of the catalyst.

To test these design ideas, we examined a series of stoichiometric reactions with acylnitrene precursors (Fig. 2A), using the iridacycle **I** and phenyl-1,4,2-dioxazole derivatives as reactants. According to the proposed reaction mechanism, the Ir-nitrenoid intermediate should be formed, and if an appropriate C–H bond were offered in close proximity, an insertion might occur more readily than the unproductive decomposition. Thus, dioxazoles bearing ortho-isopropyl substituents were presumed to be ideal, as the weak tertiary benzylic C–H bond of the isopropyl group may readily undergo C–H insertion. When (*o*-isopropyl)phenyldioxazole **1** was treated with a mixture of iridium species **I** and $NaBAr^F_4$, the Ir-dioxazole adduct **III** was formed quantitatively, as confirmed by nuclear magnetic resonance (NMR) and x-ray diffraction analysis. At slightly elevated temperature (50°C) in the presence of excess nitrile, **III** was fully converted to the C–H insertion product isoindolinone **3** with the concomitant release of a cationic iridacycle **II** and molecular acetone. With analogous 1,4,2-dioxazol-5-one (**2**), which was previously found to be a much more reactive substrate in other C–N forming reactions (*19*), the identical lactam **3** was rapidly produced even at room temperature in 5 min with CO_2 extrusion. More important, a catalytic amount of **II** was found to mediate this C–H insertion with excellent reactivity (see fig. S4), and the formation of the lactam **3** strongly supports our proposal that the postulated Ir-nitrenoid species is indeed the active intermediate and is capable of rapidly activating the relatively weak C–H bond while the undesired Curtius decomposition pathway is suppressed effectively. In contrast, when a dioxazolone bearing a flexible aliphatic chain (**4**) was used, a distinctively different result was obtained under the same conditions, leading to a six-membered Ir(III)-amido species **IV**. Although the substrate has two benzylic C–H bonds at the γ-position, the C–H insertion pathway was completely suppressed, thus highlighting that the C–N coupling and C–H insertion pathways are in competition and that the chemoselectivity is dependent on the choice of the dioxazolone substrate.

To better understand the reactivity of the Ir-nitrenoid species for the subsequent optimization of the catalysis, we carried out computer simulations on the three most plausible reaction pathways:

Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea, and Center for Catalytic Hydrocarbon Functionalizations, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea.
*These authors contributed equally to this work.
†Corresponding author. Email: sbchang@kaist.ac.kr (S.C.); mbaik2805@kaist.ac.kr (M.-H.B.)

C–N coupling, Curtius-type rearrangement, and C–H insertion (Fig. 2B). Starting from the Ir(V)-imido intermediate **VI**, transition state calculations revealed that the C–N coupling is most facile, with a barrier of only 9.1 kcal/mol. Barriers for the Curtius-type rearrangement and C–H insertion were slightly higher in energy: 9.6 kcal/mol for **VI-TS2** and 12.0 kcal/mol for **VI-TS3**, respectively. These results encouraged us to devise further strategies to avoid side reactions. As shown in Fig. 2C, we hypothesized that replacing the phenylpyridine ligand with monoanionic LX-donor ligands (X = O or N) might shut down the inner-sphere X–N coupling: Because N–N bonds [average bond dissociation energy (BDE) of ~39 kcal/mol] and O–N bonds (~50 kcal/mol) are weaker than the C–N bond (BDE of ~73 kcal/mol) (*23*), the new ligands would reduce the thermodynamic driving force for this undesired pathway. Inhibiting the Curtius-type degradation pathway in favor of the desired amidation was much more challenging. Our detailed analysis of the computer simulations revealed one potentially exploitable feature: The partial charge of the iridium center in the three transition states is remarkably different (Fig. 2B). Whereas our calculations assign a partial charge of 0.40 to the metal in the intermediate **VI**, remarkable reduction of that positive partial charge is seen for the Curtius-type rearrangement (**VI-TS2**, natural bond orbital NBO_{Ir} = 0.27). The C–H insertion is not accompanied by any notable change of partial charge, with a metal charge of 0.39 in the transition state **VI-TS3**. Hence, we can conclude that the Curtius rearrangement transition state should be much more sensitive to changes of the partial charge of the metal center than the C–H insertion, and that electron-donating ligands may increase the Curtius-rearrangement barrier to a larger extent than the C–H insertion barrier.

On the basis of this computer-aided design idea, a series of new iridium catalysts were synthesized; the catalytic reactivities of selected complexes are summarized in Fig. 2D (*24*). All iridium complexes were easy to prepare, stable in air, and convenient to handle without special precautions. To our delight, alkoxypyridyl iridium complex **VII**, which is a known precatalyst for water-oxidation chemistry (*25*), displayed some catalytic activity for the desired amination: A mixture of the lactam ***5*** and isocyanate ***6*** was obtained in 29% and 33% yield, respectively, while no N–O coupled product was observed. This result suggested that LX-type ligands can indeed prevent undesired ligand deconstruction and mediate C–H insertion chemistry. At the same time, however, formation of a large amount of isocyanate indicated that extensive ligand manipulation was necessary. The use of another N,O-chelating ligand, 8-alkoxyquinoline (**VIII**), gave rise to a notable improvement in the reaction efficiency, furnishing 73% of ***5*** at 40°C. Because the nature of the contact atom of the ligand may alter the electronic property of the metal center, we paid special attention to the influence of the X-chelating ligand component and envisioned the use of an amino group in lieu of the alkoxide. Whereas the reaction was sluggish when the alkoxy moiety was substituted with tosylamide group (**IX**), *N*-acetyl–protected aminoquinoline (**X**) showed a similar reactivity and selectivity even at room temperature. Although the *tert*-butyloxycarbonyl (Boc) group resulted in only a slight improvement (**XI**), a methyloxycarbonyl group (**XII**) further suppressed Curtius-type decomposition. At last, methoxy substitution on the ligand (5-OMe: **XIII** and 4-OMe: **XIV**) exclusively gave ***5*** in excellent yield within 6 hours and 2 hours at room temperature, respectively. This result is fully consistent with our assumption that electron-donating groups would facilitate the C–H insertion product while suppressing the formation of isocyanate. The highly active nature of catalyst **XIV** enabled a gram-scalable synthesis, and the product ***5*** was obtained in excellent yield using only 1 mole percent (mol %) of the catalyst (see fig. S1). Well-known catalytic systems for related C–H amidation, such as Ru(II)-porphyrin and dirhodium(II) complexes, were completely ineffective for the production of the lactam scaffold (*7*, *26*, *27*). These observations highlight that the iridium catalyst is particularly effective in leveraging the reactivity of acylnitrene. The present system does not require in situ protection of newly formed lactam N–H bonds to maintain catalyst activity, whereas a number of related known catalyst systems suffered from

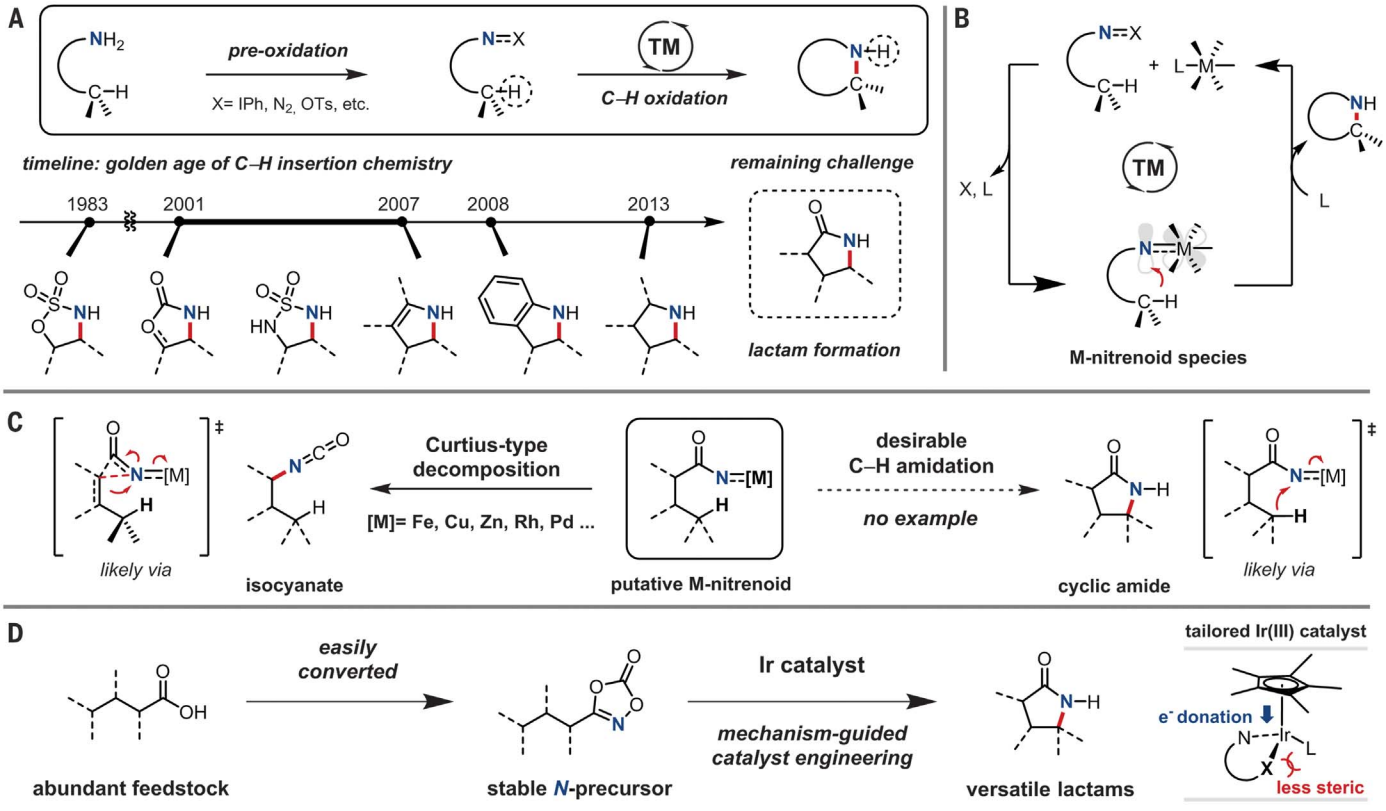


Fig. 1. A strategy for applying carbonylnitrene precursors in catalytic C–H amidation. (**A**) Examples of intramolecular C–H amination. (**B**) General catalytic mechanism. (**C**) Competitive decomposition pathway from the reactive intermediate. (**D**) Our approach with rational design of Cp*Ir(III) catalyst for γ-lactam synthesis.

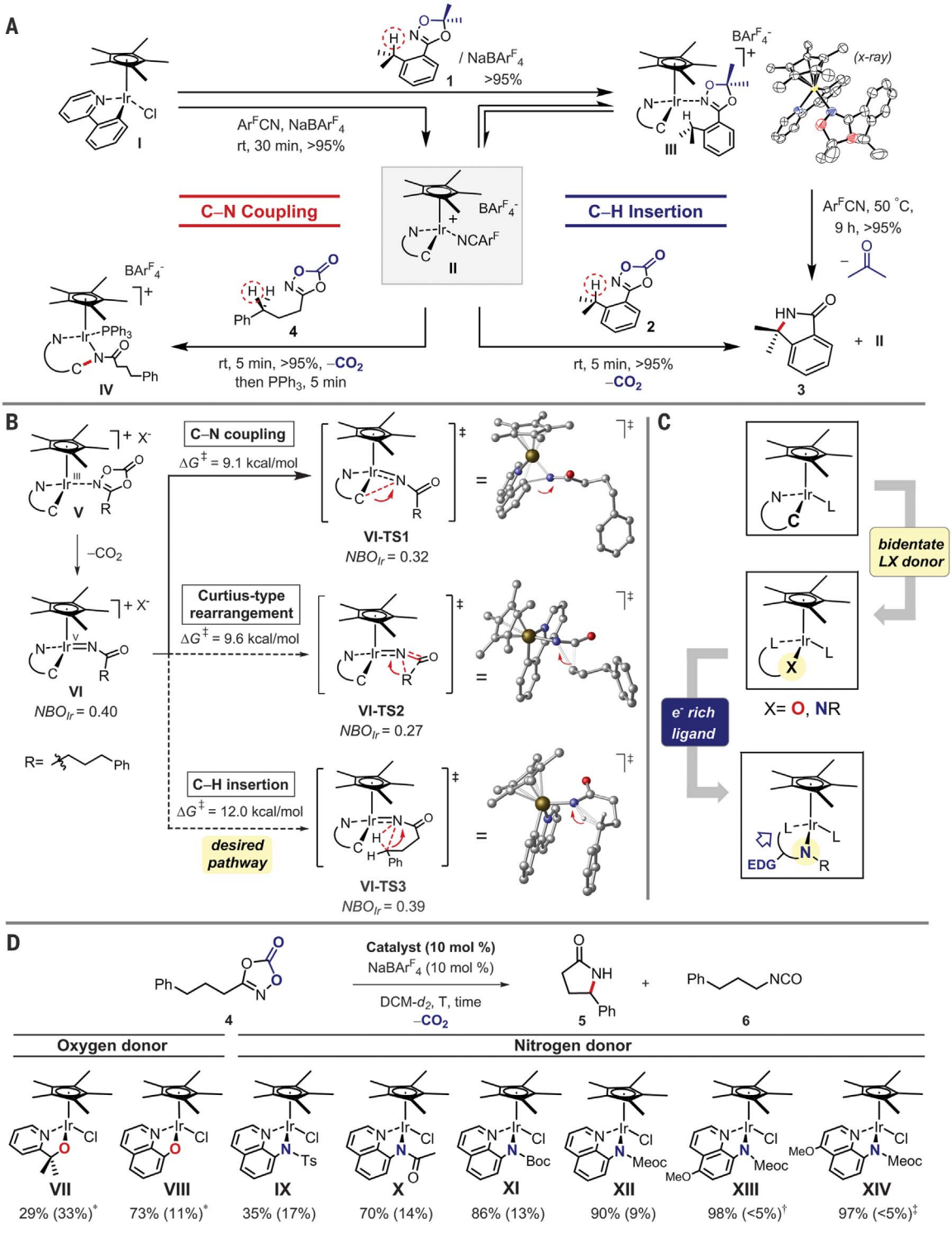


Fig. 2. Initial reaction development. (**A**) Stoichiometric studies: Viability of cyclometallated Cp*Ir(III) complex as a C–H insertion catalyst. (**B**) DFT-calculated competitive working modes from Ir(V)-acylimido species **VI**. (**C**) Principle of the catalyst design. (**D**) Catalyst optimization study. Unless otherwise indicated, reactions were run with 10 mol % of catalyst and NaBAr$^F_4$ in dichloromethane-d_2 at room temperature for 12 hours; the product yields were measured with ^{1}H NMR spectroscopy. Yields for side product **6** are indicated in parentheses. *Run at 40°C. †Run for 6 hours. ‡Run for 2 hours. Ar^F, 3,5-bis(trifluoromethyl) phenyl; rt, room temperature; Ts, *p*-toluenesulfonyl; Meoc, methyloxycarbonyl.

the deleterious inhibitory effect caused by strong coordination of the product to the catalyst (*11*, *13*), such that post-modification was necessary for catalytic turnover. Our methodology enables the direct preparation of unprotected γ-lactam products without any additional transformation.

Having identified the optimal catalysts, we investigated the generality of the iridium-catalyzed C–H amidation by exploring a wide range of substrates (Fig. 3). Various 1,4,2-dioxazol-5-ones were accessed from abundant carboxylic acid feedstock by two-step sequences consisting of acid activation/hydroxamic acid formation followed by carbonylative cyclization; both of these steps are highly efficient, and the dioxazolones can be easily prepared in excellent yields. The cyclization of substrates containing benzylic C–H bonds with catalyst **XIV** provided the corresponding γ-lactams in high yields (**5**, **7**–**10**), even in the presence of a Boc-protected free N–H group (**11**). The formation of 2,2-dimethyl lactam **12**, however, required slightly higher temperature (80°C), presumably as a result of the steric crowding adjacent to the dioxazolone moiety. Functionalization of a secondary benzylic C–H bond in the cyclopentyl ring resulted in *cis*-tricyclic γ-lactam (**13**) as a sole

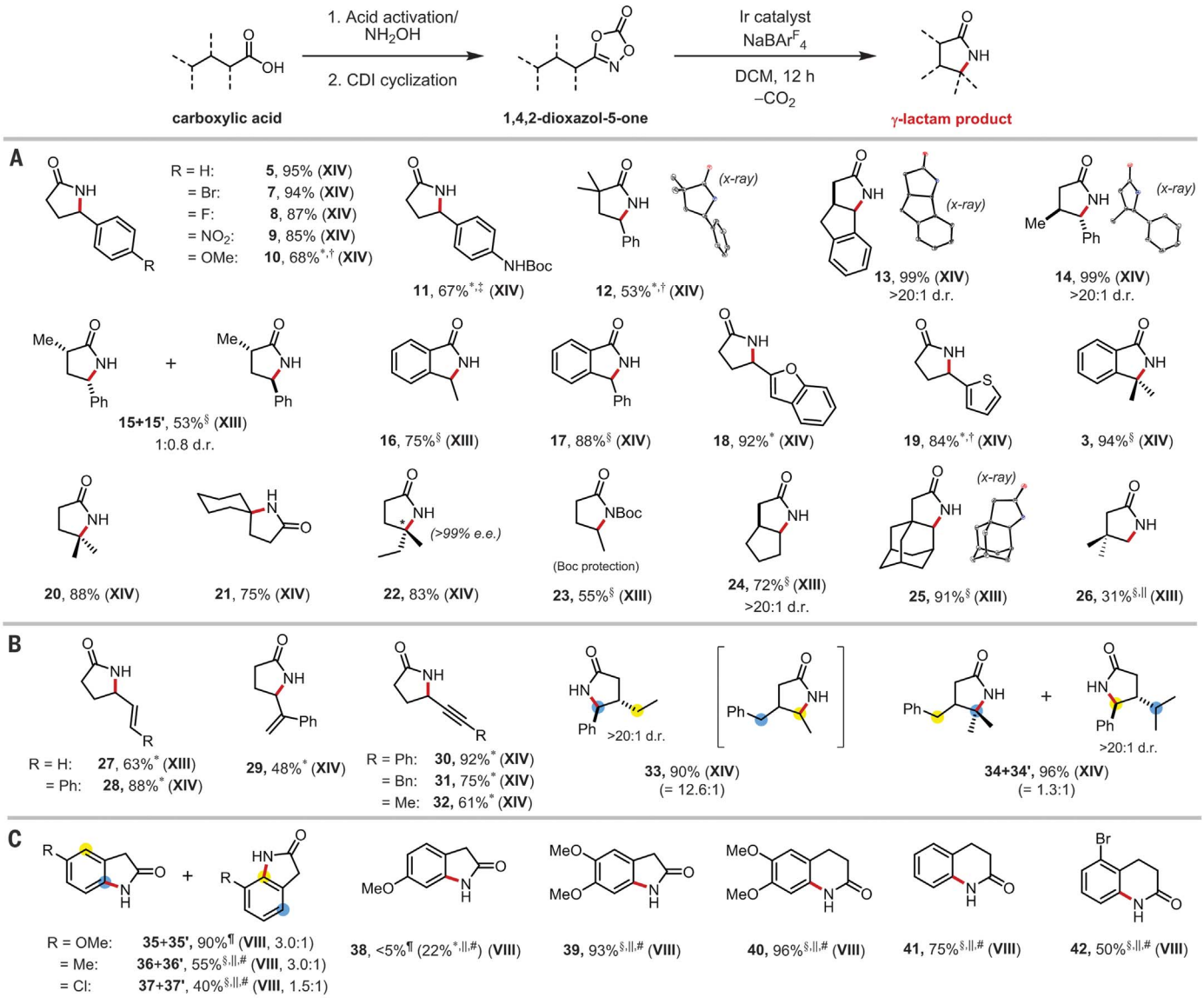


Fig. 3. Reaction scope of Ir(III)-catalyzed intramolecular amidation. (**A**) Functionalization of benzylic, tertiary, secondary, and primary C(sp^3)–H bonds. (**B**) Examination of chemo- and regioselectivity. (**C**) Functionalization of aromatic C(sp^2)–H bonds. Unless otherwise indicated, reactions were run with 2 mol % of catalyst and $NaBAr^F_4$ at 40°C for 12 hours. The Ir catalyst used is indicated in parentheses; the data are reported as percent isolated yields. *10 mol % catalyst was used. †Run at 80°C for 48 hours. ‡Run at 40°C for 12 hours followed by 80°C for 24 hours. §5 mol % catalyst was used. ||Hexafluoro-2-propanol was used as solvent. ¶Run at room temperature. #Run at 60°C. d.r., diastereomeric ratio; e.e., enantiomeric excess.

product in 99% yield. Whereas a β-methyl substituent gave rise to excellent reactivity and diastereoselectivity (**14**, >20:1), the amination of an α-substituted substrate gave a mixture of syn- and anti-inserted products without a bias (**15**, **15′**, 1:0.8). When dioxazolones prepared from benzoic acids were used, 3-substituted isoindolin-1-ones (**16**, **17**) were obtained in high yields. Heterocycles such as benzofuran (**18**) or thiophene (**19**) were also tolerated under the reaction conditions. The method also transformed tertiary C–H bonds, thereby introducing quaternary carbon centers (**3**, **20**) and an azaspirocyclic scaffold (**21**). The desired lactamization proceeded in a stereospecific manner when γ-chiral dioxazolone was used (**22**), providing a clue to the insertion mechanism (see below). Finally, amidation of nonactivated secondary C–H bonds of *n*-butyl (**23**), cyclopentyl (**24**), and adamantyl (**25**) groups was successful, and primary C(sp^3)–H bonds also underwent reaction, albeit in moderate yield (**26**).

The high reactivity of the Ir catalyst motivated us to examine the selectivity trends when multiple reactive sites were embedded in the same molecule (Fig. 3B). Preferential chemo- and regioselectivity patterns often provide clues to the reaction mechanism and have been widely used to understand microscopic insertion pathways (*28*). When γ-allylic C–H bonds that potentially undergo olefin aziridination were tested, insertion products were observed exclusively with no detectable sign of aziridine formation (**27**–**29**). Olefin isomerization was not observed during the formation of lactam **28**. Similarly, dioxazolones bearing γ-propargylic C–H bonds underwent lactamization in high yields (**30**–**32**), allowing for efficient access to key precursors in the total synthesis of (–)-stemoamide derivatives (*29*). Considering that intramolecular amidation has frequently suffered from related aziridination reactions caused by electron-rich π functionality (*28*, *30*, *31*), the remarkable chemoselectivity observed in this system is intriguing (*32*). Next, we designed substrates having nonequivalent γ and γ′ C–H bonds to qualitatively rank the reactivity toward distinctive C–H bonds. Benzylic C–H bonds were favored over nonactivated secondary C–H bonds (**33**, 12.6:1) but were

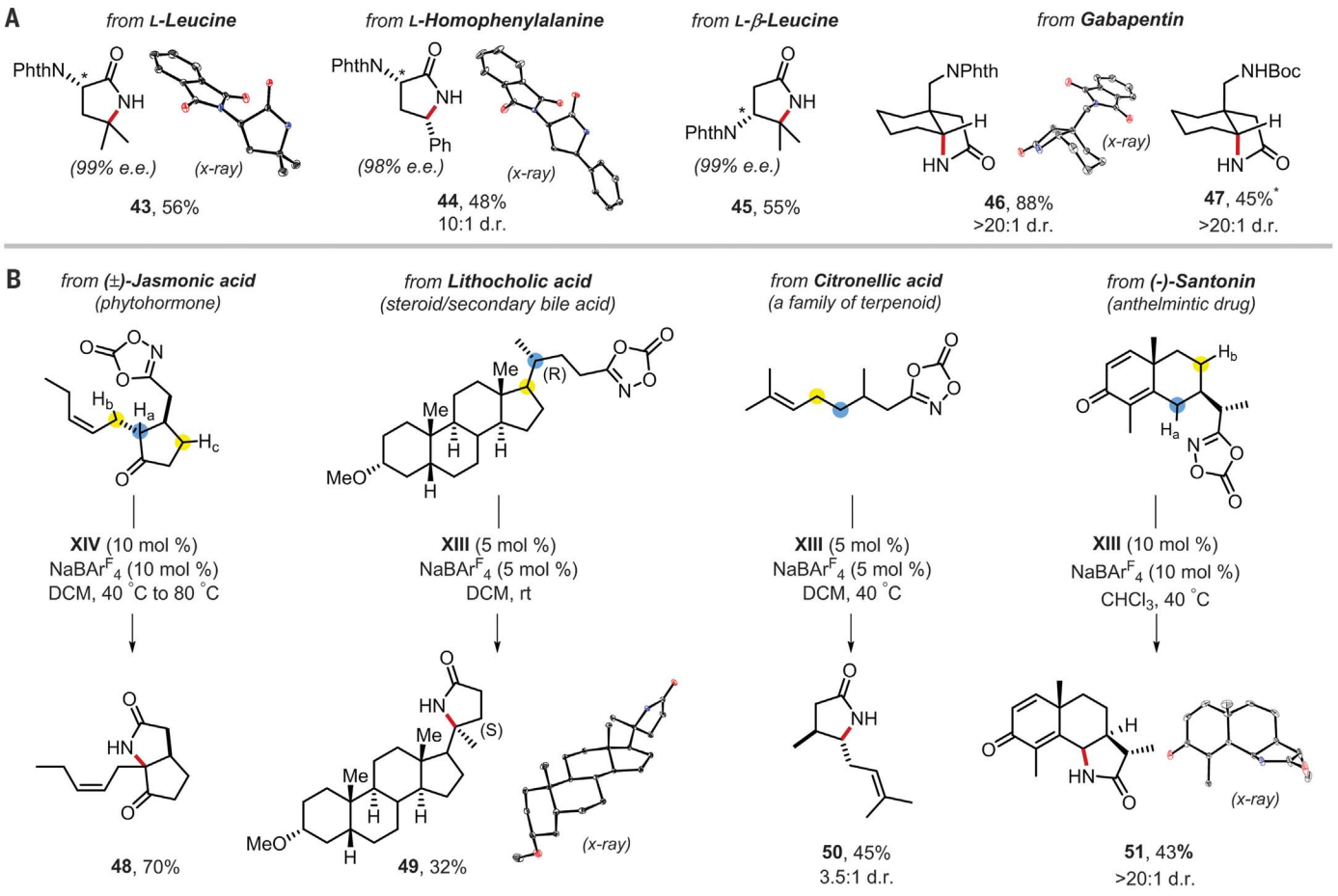


Fig. 4. Synthetic utility of the intramolecular C–H amidation. (**A**) Derivatization of amino acids to versatile γ-lactams. (**B**) Late-stage functionalization of complex molecules. Unless otherwise indicated, reactions were run with 5 mol % of catalyst **XIII** and $NaBAr^F_4$ at 40°C for 12 hours. *Hexafluoro-2-propanol was used as solvent.

disfavored relative to tertiary C–H bonds (**34**, **34′**, 1.3:1). The observed reactivity is not matched qualitatively with trends in homolytic bond cleavages, as the BDE of a secondary benzylic C–H bond (86 kcal/mol) is lower than that of a tertiary C–H bond (96 kcal/mol) (*33*). This finding implies that radical-type hydrogen atom abstraction may not be operative (*8*, *30*, *31*). In this context, intramolecular kinetic isotope experiment (KIE) results suggest a concerted C–H insertion mechanism. When deuterated dioxazolone (**4**-d_2) was subjected to the amidation, the intramolecular KIE value was found to be 1.51 ± 0.01 (table S3). This value is similar to the KIE obtained in the related Rh_2(II)-catalyzed amidation of sulfamates (*28*) and is lower than the values reported typically for H-atom abstraction mechanisms (*13*, *34*). In addition, the stereospecificity of the amination, which was demonstrated by the chiral lactam **22** in Fig. 3A, corroborated the concerted mechanism, although fast radical rebound cannot be ruled out. Density functional theory (DFT) calculations of the proposed mechanisms further indicated that the concerted insertion pathway from the singlet Ir-nitrenoid species is energetically more feasible than the involvement of a triplet nitrenoid traversing a radical abstraction pathway (reaction barrier $\Delta G^{\ddagger}$ = 6.5 and 14.2 kcal/mol, respectively). Detailed mechanistic descriptions including reaction energy profiles are shown in fig. S8.

We next extended the method to amidation of aromatic C(sp^2)–H bonds: 1,4,2-dioxazol-5-ones derived from 2-phenylacetic acids were successfully converted to the corresponding indolinones using catalyst **VIII** (Fig. 3C). Specifically, constitutional isomers of indolinone products **35** and **35′** were obtained in high yields with 3:1 regioselectivity, presumably due to steric hindrance at the meta position. Other meta-substituents such as methyl (**36**) and chloro (**37**) groups were also smoothly incorporated with similar selectivity. In contrast, a methoxy substituent in the para position significantly diminished reactivity (**38**), implying that an electrophilic aromatic substitution-type pathway might be operative (*14*). 3,4-Dimethoxy-bearing dioxazolone gave rise to a single product **39** in excellent yield. Six-membered dihydroquinolinones **40**–**42** were also produced by the same procedure. Catalysts **XIII** and **XIV**, which were efficient in the C(sp^3)–H amidation, were less active in this sp^2 C–H amidation, which suggests that the reactions have slightly different electronic demands (*24*).

Because 1,4,2-dioxazol-5-ones can be conveniently prepared in high yields from abundant carboxylic acids, we anticipated that the present amidation protocol would provide an efficient means of preparing valuable γ-lactams including biorelevant molecules. To our delight, a wide range of amino acids were readily converted to γ-lactams under the optimal conditions (Fig. 4A). For instance, the reaction of *N*-phthalimide–protected L-leucine derivatives provided the corresponding cyclic amide **43** in 56% yield without epimerization at the α-carbon center (99% enantiomeric excess). When a substrate derived from unnatural L-homophenylalanine was used, syn-addition of the amino group was found to occur more favorably over the formation of anti-inserted products (**44**, syn/anti = 10:1). Not only α- but also β- and γ-amino acids, such as L-β-leucine and gabapentin derivatives, were successfully converted to the corresponding lactams **45**–**47** in acceptable yields, respectively.

The amidation protocol was also highly effective in more complex molecules that possessed multiple reactive C–H bonds (Fig. 4B) (*35*). In a

reaction of a jasmonic acid derivative having three potentially reactive sites (tertiary, allylic, and secondary C–H bonds), tertiary C–H inserted product **48** was observed exclusively without any other detectable regioisomers. To further test the applicability of this method for late-stage functionalizations, we examined a family of bile acid and terpenoid substrates. Dioxazolones derived from lithocholic acid underwent the intramolecular amidation, leading to a steroid-type lactam scaffold (**49**). In addition, the amide moiety was selectively formed at the secondary γ-C–H bond of a citronellic acid derivative (**50**) even in the presence of activated C–H bonds at the δ-position. A lactam surrogate of (–)-α-santonin (**51**) was successfully synthesized under the optimal conditions, and the C–N bond formation took place selectively at the allylic position.

Our findings show that a mechanism-guided approach can be used to resolve a long-standing challenge in catalytic C–H amidation chemistry. Although this work falls short of definitively proving the existence of the Ir(V)-nitrenoid intermediate, all of the data and insights gained are fully consistent and strongly support its putative role.

REFERENCES AND NOTES

1. H. Kim, S. Chang, *Acc. Chem. Res.* **50**, 482–486 (2017).
2. H. M. L. Davies, J. R. Manning, *Nature* **451**, 417–424 (2008).
3. G. Dequirez, V. Pons, P. Dauban, *Angew. Chem. Int. Ed.* **51**, 7384–7395 (2012).
4. Y. Park, Y. Kim, S. Chang, *Chem. Rev.* **117**, 9247–9301 (2017).
5. R. Breslow, S. H. Gellman, *J. Am. Chem. Soc.* **105**, 6728–6729 (1983).
6. X.-Q. Yu, J.-S. Huang, X.-G. Zhou, C.-M. Che, *Org. Lett.* **2**, 2233–2236 (2000).
7. J. L. Roizen, M. E. Harvey, J. Du Bois, *Acc. Chem. Res.* **45**, 911–922 (2012).
8. S. M. Paradine *et al*., *Nat. Chem.* **7**, 987–994 (2015).
9. C. K. Prier, R. K. Zhang, A. R. Buller, S. Brinkmann-Chen, F. H. Arnold, *Nat. Chem.* **9**, 629–634 (2017).
10. M. Shen, B. E. Leslie, T. G. Driver, *Angew. Chem. Int. Ed.* **47**, 5056–5059 (2008).
11. Q. Nguyen, K. Sun, T. G. Driver, *J. Am. Chem. Soc.* **134**, 7262–7265 (2012).
12. B. J. Stokes, H. Dong, B. E. Leslie, A. L. Pumphrey, T. G. Driver, *J. Am. Chem. Soc.* **129**, 7500–7501 (2007).
13. E. T. Hennessy, T. A. Betley, *Science* **340**, 591–595 (2013).
14. M. P. Paudyal *et al*., *Science* **353**, 1144–1147 (2016).
15. E. F. V. Scriven, K. Turnbull, *Chem. Rev.* **88**, 297–368 (1988).
16. H. Lebel, O. Leogane, *Org. Lett.* **7**, 4107–4110 (2005).
17. D. Li, T. Wu, K. Liang, C. Xia, *Org. Lett.* **18**, 2228–2231 (2016).
18. S. Bräse, C. Gil, K. Knepper, V. Zimmermann, *Angew. Chem. Int. Ed.* **44**, 5188–5240 (2005).
19. Y. Park, K. T. Park, J. G. Kim, S. Chang, *J. Am. Chem. Soc.* **137**, 4534–4542 (2015).
20. D. Willcox *et al*., *Science* **354**, 851–857 (2016).
21. J. Ryu, J. Kwak, K. Shin, D. Lee, S. Chang, *J. Am. Chem. Soc.* **135**, 12861–12868 (2013).
22. Y. Park, J. Heo, M.-H. Baik, S. Chang, *J. Am. Chem. Soc.* **138**, 14020–14029 (2016).
23. P. Atkins, L. Jones, in *Chemical Principles* (Freeman, ed. 5, 2010), pp. 80–81.
24. See the supplementary materials for a full list of tested catalysts for the optimization.
25. N. D. Schley *et al*., *J. Am. Chem. Soc.* **133**, 10473–10481 (2011).
26. V. Bizet, L. Buglioni, C. Bolm, *Angew. Chem. Int. Ed.* **53**, 5639–5642 (2014).
27. Treatment of **4** with Ru(TPP)CO catalyst mainly resulted in the Curtius-type decomposition products, whereas **4** was quantitatively recovered with $Rh_2(OAc)_4$ and $Rh_2(esp)_2$ catalysts (TPP = tetraphenylporphyrinato; esp = α,α,α′,α′-tetramethyl-1,3-benzenedipropionate). See table S1 and the associated description for details.
28. K. W. Fiori, C. G. Espino, B. H. Brodsky, J. Du Bois, *Tetrahedron* **65**, 3042–3051 (2009).
29. P. A. Jacobi, K. Lee, *J. Am. Chem. Soc.* **122**, 4295–4303 (2000).
30. M. E. Harvey, D. G. Musaev, J. Du Bois, *J. Am. Chem. Soc.* **133**, 17207–17216 (2011).
31. J. M. Alderson, A. M. Phelps, R. J. Scamp, N. S. Dolan, J. M. Schomaker, *J. Am. Chem. Soc.* **136**, 16720–16723 (2014).
32. When a substrate that was shortened by one carbon between the dioxazolone and double-bond moiety was treated with **XIV**, an Ir-amido complex was formed quantitatively, plausibly via olefin aziridination followed by ring-opening processes. See figs. S10 and S11 for details.
33. Y.-R. Luo, *Handbook of Bond Dissociation Energies in Organic Compounds* (CRC Press, 2002).
34. S.-M. Au, J.-S. Huang, W.-Y. Yu, W.-H. Fung, C.-M. Che, *J. Am. Chem. Soc.* **121**, 9120–9132 (1999).
35. Z. Huang, G. Dong, *Acc. Chem. Res.* **50**, 465–471 (2017).

ACKNOWLEDGMENTS

Funding: Supported by Institute for Basic Science (Republic of Korea) grant IBS-R010-D1. **Author contributions:** S.Y.H., Y.P., M.-H.B., and S.C. conceived and designed the project and wrote the manuscript; S.Y.H., Y.P., Y.H., and Y.B.K. carried out the experiments; S.Y.H. and Y.P. performed DFT calculations; S.C. organized the research; and all authors analyzed the data, discussed the results, and commented on the manuscript. **Competing interests:** S.Y.H., Y.P., Y.H., Y.B.K., and S.C. are inventors on patent application numbers KR10-2018-0000421 and KR10-2018-0000449, submitted by IBS and KAIST, that cover preparation and application of the related transition metal catalysts. **Data and materials availability:** The supplementary materials contain computational details, NMR spectra, and HPLC traces. Crystallographic data are available free of charge from the Cambridge Crystallographic Data Centre under reference numbers CCDC 1569045 to CCDC 1569058 and CCDC 1587791.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1016/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S19
NMR Spectra
HPLC Traces
References (*36*–*80*)

22 August 2017; resubmitted 1 December 2017
Accepted 5 January 2018
10.1126/science.aap7503

GEOLOGY

Evolution of alluvial mudrock forced by early land plants

William J. McMahon and Neil S. Davies*

Mudrocks are a primary archive of Earth's history from the Archean eon to recent times, and their source-to-sink production and deposition play a central role in long-term ocean chemistry and climate regulation. Using original and published stratigraphic data from all 704 of Earth's known alluvial formations from the Archean eon (3.5 billion years ago) to the Carboniferous period (0.3 billion years ago), we prove contentions of an upsurge in the proportion of mud retained on land coeval with vegetation evolution. We constrain the onset of the upsurge to the Ordovician-Silurian and show that alluvium deposited after land plant evolution contains a proportion of mudrock that is, on average, 1.4 orders of magnitude greater than the proportion contained in alluvium from the preceding 90% of Earth's history. We attribute this shift to the ways in which vegetation revolutionized mud production and sediment flux from continental interiors.

Earth's stratigraphic record preserves a number of trends in biogenic and chemogenic sedimentary rocks through time, reflecting secular changes at the surface of the planet (*1*). Siliciclastic sediments, produced primarily by the mechanical and chemical breakdown of parent rock, do not have such first-order biological controls. However, subtle secular changes have been previously quantified, including both clay mineral evolution (*2*) and changes in (bio)geomorphic sedimentary structures and architecture (*3*–*5*). In terms of bulk lithology, it is a long-held anecdotal contention (*4*, *6*–*8*) that mudrock is rare in alluvium that was deposited before the evolution of land plants but is common in alluvium deposited thereafter. We quantitatively tested this contention and found it to be true, demonstrating the magnitude and timing of the onset of the increase using data recording the proportional thickness of mudrock within alluvial stratigraphic sections (Fig. 1A).

We surveyed 1196 published reports and undertook 125 original field investigations to gather data on Earth's 704 known, globally distributed Archean-Carboniferous alluvial stratigraphic units and compiled a single database (table S1). Data reduction and analysis show that mudrock is a negligible component of alluvial strata deposited during the first ~3.0 billion years of Earth's sedimentary rock record but is common or dominant after the middle Paleozoic [mudrock is defined lithologically as all rocks dominantly composed of detrital and weathered sedimentary grains of ≤0.063 mm (siltstone) (*9*)]. In Archean strata [4000 to 2500 million years (Ma) old], the cumulative stratigraphic proportion of mudrock within alluvial strata ranges between 0 and 14% (median, 1.0%), whereas in Carboniferous strata (359 to 299 Ma old) the range is 0 to 90% (median, 26.2%) (Fig. 1D). LOESS regression analysis of the data constrains the upsurge between the Late Ordovician and the Devonian (458 to 359 Ma ago) (Fig. 1B), after which the range and average proportion of mudrock in alluvium never reverted to the same low values that characterized the first 3 billion years (Ga) of Earth's stratigraphic record. Subsampling of the data shows that the amount of mudrock was 1.1 orders of magnitude greater in the Late Ordovician to the Silurian (458 to 419 Ma ago), 1.3 orders of magnitude greater in the Early to Middle Devonian (418 to 379 Ma ago), 1.45 orders of magnitude greater in the Late Devonian to the early Carboniferous (378 to 339 Ma ago), and 1.75 orders of magnitude greater in the middle to late Carboniferous (338 to 299 Ma ago) than in the Archean to the Middle Ordovician (3500 to 458 Ma ago) (Fig. 1C).

This stratigraphically unidirectional upsurge in alluvial mudrock likely rules out a cause due to episodic or cyclic geological phenomena (such as tectonic or climatic controls) that persisted on Earth throughout the Archean to the Carboniferous (*10*, *11*) (fig. S11). The first 3 Ga of the interval we studied included multiple alternations between icehouse and greenhouse conditions (*12*), the assembly of at least two supercontinents (*13*), and 16 known regional orogenies (*14*). None of these events had any apparent influence on the near-uniform global scarcity of preserved alluvial mudrock. Similarly, the onset of the trend does not correlate with other prominent potential triggers in the geological record. For example, it postdates Paleoproterozoic oxygenation by at least 1640 Ma (*15*), Neoproterozoic oxygenation by 142 Ma (*15*), and the advent of microbial life on land by 2540 Ma (*16*) and may predate the increased survivorship of nonmarine strata by up to 60 Ma (*11*, *17*). The systematic misidentification of pre-Ordovician mudrock as marine in previous studies is a potential source of uncontrolled bias in our study (*11*). However, after testing the data against various alternative hypotheses (*11*) (figs. S5 to S8 and S11 to S13), we argue that the most plausible explanation is that prevegetation Earth had distinct syndepositional controls on sedimentation that discouraged the production or accumulation of alluvial mudrock. The trend mirrors the fossil plant record (*18*–*20*), and the appearance of primitive plants would have introduced three mechanisms important for producing mudrock-rich alluvial strata. Plants lead to an increased production of the directly weathered fraction of fines (clays) (*2*, *18*, *21*–*26*). They also increase retention of all (weathered and detrital) fines in continental deposystems through binding (the fastening of masses of grains by plant parts such as roots) (*25*, *26*). Finally, the process of baffling (the capture and forced deposition of grains from within a moving fluid passing over and around plant parts) also increases retention of all (weathered and detrital) fines in continental deposystems (*27*, *28*).

Terrigenous fines are sourced into sedimentary systems through the mechanical mass wasting of chemical weathering profiles, supplying both weathered and detrital silt, mud, and clay particles (*21*). Land plants are not a prerequisite for the mechanical production of fines, and abiotic, microbial, and fungal processes could all promote the silicate weathering of clays on prevegetation Earth (*16*, *18*, *21*, *29*, *30*). The presence of minor mudrock in alluvium of all ages, in addition to known terrigenous mudrocks from prevegetation lacustrine and marine facies, demonstrates these alternative pathways (Fig. 1A). However, land plants do promote the production of clay minerals and the depth of chemical weathering profiles by increasing atmosphere-substrate connectivity through rooting, through the direct secretion of organic acids and chelates, and by developing symbiotic relationships that increase the capacity of cyanobacteria and fungi to dissolve soil grains (*2*, *18*, *21*–*26*). The degree to which the earliest bryophyte-grade plants could have boosted silicate weathering (*16*, *22*, *23*, *31*, *32*) remains a point of debate, but a clear global intensification followed the evolution of a deeper-rooted Devonian flora (*18*, *22*, *24*, *25*). The initial range expansion of mudrock proportions in the Ordovician-Silurian (Fig. 1B) suggests that even the earliest plants played some role in promoting mudrock in alluvium (*26*), before the pronounced rise seen after the Devonian evolution of rooting. However, even if the earliest bryophytes increased weathering, net production alone may not account for the trend. In limited instances where mudrock type has previously been distinguished, siltstone abundance exhibits the same unidirectional trend as mudstone, claystone, and shale abundance (*11*) (fig. S2), suggesting that even fines with a greater (though not exclusive) probability of having been mechanically and abiotically generated (*21*) are diminished in prevegetation alluvium.

Before vegetation, continents were colonized by microbial mats (*16*), but the lack of below-ground structure to these communities meant

Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK.
*Corresponding author. Email: nsd27@cam.ac.uk

that they were prone to undercutting and reworking by fluvial channels and so had a negligible effect on the retention of sediment (*33*). In contrast, the establishment of root systems offered novel mechanical protection against the erosion of sediment in alluvial settings (*23*, *25*) and would thus have promoted the physical retention of clay, mud, and silt. This importance of below-ground stabilization would clearly have played some role in the major Devonian upsurge in mudrock, but the root systems of earlier land plants were limited (*18*, *23*), so this is an unlikely explanation for observed mud-rich Ordovician and Silurian formations.

The above-ground structures of even shallow-rooted and small-stature vegetation today can reduce near-surface flow of water and wind below a critical velocity that promotes sediment deposition (*27*, *28*). Observations of mosses and liverworts show effective trapping of individual fine grains between their stems, leaves, and thalli, incorporating sediment into cryptogamic ground covers (*26*). Even though direct physiological analogy between modern and early land plants may be inappropriate (*18*), the earliest above-ground plant structures must have introduced a wholly unprecedented biological component of roughness to Earth's surface. This suggests a large role for baffling by even primitive above-ground plant constructions, promoting the recurrence frequency of deposition of fines in the alluvial realm and contributing to the mudrock increase.

The Paleozoic increase in alluvial mudrock is an important characteristic of the global sedimentary geological record. The timing relative to the appearance of plants is unlikely to be a coincidence, as plants can greatly contribute to the development and retention of alluvial mudrocks. The source-to-sink deposition of prevegetation mud was thus profoundly different from that seen in the present day (*34*). On prevegetation Earth, all fines had limited potential for final (preserved) deposition within continental conduits, regardless of any non–vegetation-related variations in chemical weathering intensity (*29*, *30*, *35*) or sediment flux (*36*). Archean to Middle Ordovician marine settings would have received a generally greater flux of whatever terrigenous fines were being produced in continental source areas. After the Late Ordovician, and to a greater degree after the Devonian, an increasing proportion of terrigenous fines were produced and/or retained on the continents; thus, the marine realm may have received a diminished fraction of total continentally weathered fines. This need not have equated to a diminished volume because net production at the source would have been greater. A fuller understanding of mudrock in the absence of vegetation is a prerequisite for any studies that invoke ancient terrestrial mudrock strata as a primary archive of geochemical or petrological data and will have implications for understanding the context and nature of mudrocks that are increasingly detected on non-vegetated planets such as Mars (*8*, *37*).

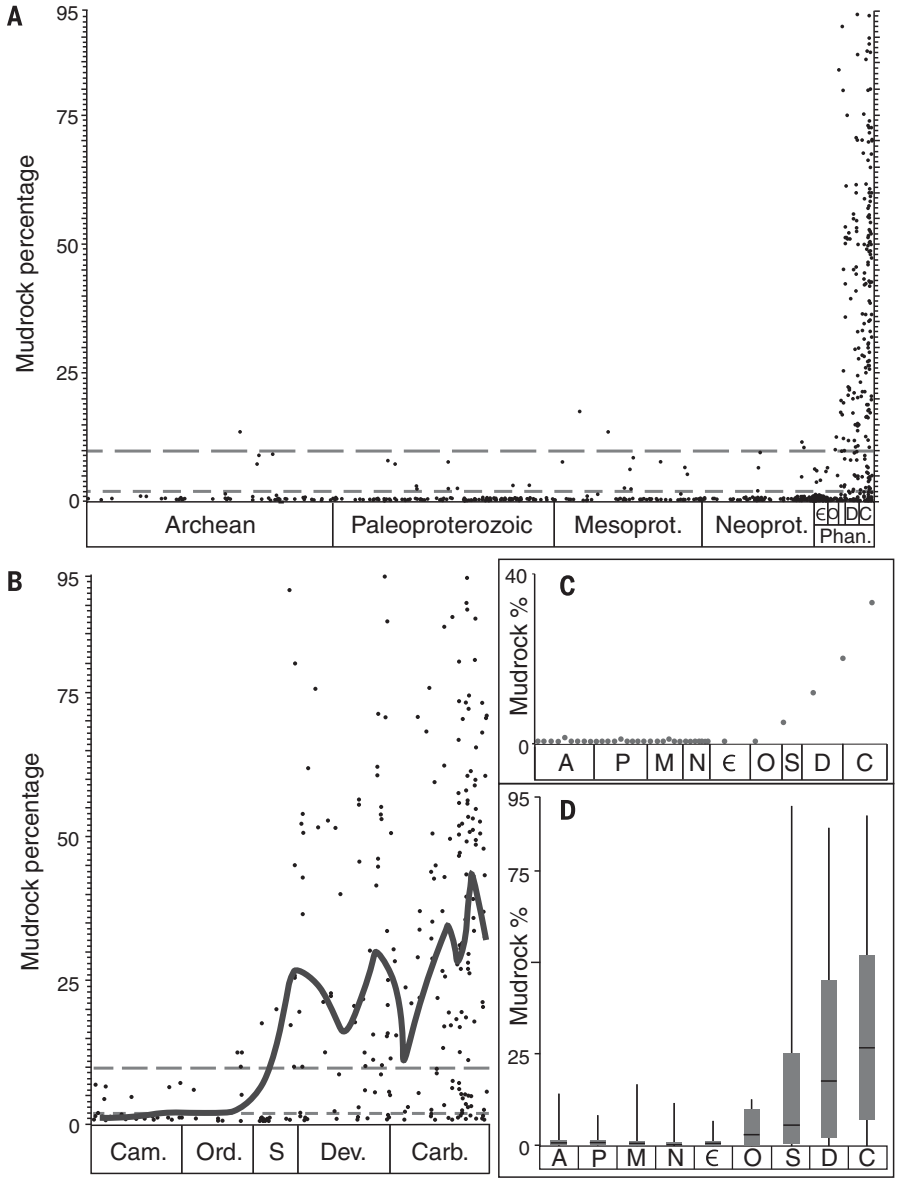


Fig. 1. The range and maximum proportion of mudrock in alluvial successions increase markedly after the evolution of vegetation. The proportion of mudrock within alluvial successions (percentage of vertical stratigraphic thickness) is plotted against geologic age [the *x* axis is scaled to numerical ages, with the start of intervals based on the Geologic Time Scale 2012 (*38*): Archean (A; 4000 Ma ago), Paleoproterozoic (P; 2500 Ma ago), Mesoproterozoic (Mesoprot. or M; 1600 Ma ago), Neoproterozoic (Neoprot. or N; 1000 Ma ago), Cambrian (∈ or Cam.; 541.0 Ma ago), Ordovician (O or Ord.; 485.4 Ma ago), Silurian (S; 443.8 Ma ago), Devonian (D or Dev.; 419.2 Ma ago), Carboniferous (C or Carb.; 358.9 Ma ago), and Permian (298.9 Ma ago)]. (**A**) Each individual point records one of the known 594 alluvial stratigraphic units deposited during this interval. Long-dashed line, 10%; short-dashed line, 2%. Phan., Phanerozoic. (**B**) Enlarged plot for the Phanerozoic with LOESS regression line (solid gray line). LOESS was conducted with a smoothing parameter of 0.9. (**C**) Proportion of mudrock corrected for variation in sampling intensity by subsampling. Each individual point represents the median value seen across 100 individual subsampling trials (see supplementary materials for methodology). (**D**) Median, range, upper quartile, and lower quartile of mudrock proportion for each interval.

REFERENCES AND NOTES

1. P. G. Eriksson *et al.*, *Gondwana Res.* **24**, 468–489 (2013).
2. R. M. Hazen *et al.*, *Am. Mineral.* **98**, 2007–2029 (2013).
3. W. E. Dietrich, J. T. Perron, *Nature* **439**, 411–418 (2006).
4. N. S. Davies, M. R. Gibling, *Earth Sci. Rev.* **98**, 171–200 (2010).
5. N. S. Davies, M. R. Gibling, *Nat. Geosci.* **4**, 629–633 (2011).

6. D. Winston, in *Memoir*, vol. 5, *Fluvial Sedimentology*, A. D. Miall, Ed. (Canadian Society of Petroleum Geologists, 1978), p. 343.
7. D. S. McCormick, J. P. Grotzinger, *J. Sediment. Petrol.* **63**, 398–419 (1993).
8. J. P. Grotzinger *et al*., *Science* **343**, 1242777 (2014).
9. A. G. Ilgen *et al*., *Earth Sci. Rev.* **166**, 132–152 (2017).
10. N. S. Davies *et al*., *J. Geol. Soc. London* **174**, 947–950 (2017).
11. Materials and methods are available as supplementary materials.
12. P. F. Hoffman, *Geol. Today* **25**, 100–107 (2009).
13. D. C. Bradley, *Earth Sci. Rev.* **108**, 16–33 (2011).
14. T. H. Torsvik, L. R. M. Cocks, *Earth History and Palaeogeography* (Cambridge Univ. Press, 2016).
15. T. W. Lyons, C. T. Reinhard, N. J. Planavsky, *Nature* **506**, 307–315 (2014).
16. T. M. Lenton, S. J. Daines, *New Phytol.* **215**, 531–537 (2017).
17. S. E. Peters, J. M. Husson, *Geology* **45**, 323–326 (2017).
18. C. K. Boyce, J.-E. Lee, *Annu. Rev. Earth Planet. Sci.* **45**, 61–87 (2017).
19. C. V. Rubinstein, P. Gerrienne, G. S. de la Puente, R. A. Astini, P. Steemans, *New Phytol.* **188**, 365–369 (2010).
20. K. K. S. Matsunaga, A. M. F. Tomescu, *Ann. Bot. (London)* **117**, 585–598 (2016).
21. H. W. Nesbitt, C. M. Fedo, G. M. Young, *J. Geol.* **105**, 173–192 (1997).
22. J. Quirk *et al*., *Biol. Lett.* **8**, 1006–1011 (2012).
23. D. Edwards, L. Cherns, J. A. Raven, *Palaeontology* **58**, 803–837 (2015).
24. J. L. Morris *et al*., *Palaeontology* **58**, 787–801 (2015).
25. J. Xue *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9451–9456 (2016).
26. R. L. Mitchell *et al*., *Geology* **44**, 1007–1010 (2016).
27. A. Gurnell, *Earth Surf. Process. Landforms* **39**, 4–25 (2014).
28. H. Moor *et al*., *J. Ecol.* **105**, 1623–1635 (2017).
29. N. J. Tosca *et al*., *Geochim. Cosmochim. Acta* **74**, 1579–1592 (2010).
30. M. Kennedy, M. Droser, L. M. Mayer, D. Pevear, D. Mrofka, *Science* **311**, 1446–1449 (2006).
31. J. Quirk *et al*., *Proc. R. Soc. London Ser. B* **282**, 20151115 (2015).
32. P. Porada *et al*., *Nat. Commun.* **7**, 12113 (2016).
33. W. J. McMahon, N. S. Davies, D. J. Went, *Precambrian Res.* **292**, 13–34 (2017).
34. E. L. Leithold, N. E. Blair, K. W. Wegmann, *Earth Sci. Rev.* **153**, 30–42 (2016).
35. P. L. Corcoran, W. U. Mueller, in *Precambrian Sedimentary Environments: A Modern Approach to Ancient Depositional Systems*, W. Altermann, P. L. Corcoran, Eds. (Blackwell, 2002), pp. 183–211.
36. S. E. Peters, R. R. Gaines, *Nature* **484**, 363–366 (2012).
37. J. Schieber *et al*., *Sedimentology* **64**, 311–358 (2017).
38. F. M. Gradstein, J. G. Ogg, F. J. Hilgen, *Newsl. Stratigr.* **45**, 171–188 (2012).

ACKNOWLEDGMENTS

We thank M. R. Gibling for assistance with Paleozoic data during the course of a previous project and O. Shorttle and J. P. Mattern for advice on the statistical treatment of the data. **Funding:** W.J.M. was funded by Shell International Exploration and Production under Research Framework agreement PT38181. **Author contributions:** The authors contributed equally to the collection and analysis of the field and literature data presented. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** A full reference list for the data reported in this paper is included in the supplementary materials (table S1).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1022/suppl/DC1
Materials and Methods
Figs. S1 to S15
Table S1
References (*39*–*71*)

17 April 2017; resubmitted 26 May 2017
Accepted 16 January 2018
10.1126/science.aan4660

NEUROSCIENCE

Evidence for a neural law of effect

Vivek R. Athalye,[1,2]* Fernando J. Santos,[1]* Jose M. Carmena,[2,3,4]†‡ Rui M. Costa[1,5]†‡

Thorndike's law of effect states that actions that lead to reinforcements tend to be repeated more often. Accordingly, neural activity patterns leading to reinforcement are also reentered more frequently. Reinforcement relies on dopaminergic activity in the ventral tegmental area (VTA), and animals shape their behavior to receive dopaminergic stimulation. Seeking evidence for a neural law of effect, we found that mice learn to reenter more frequently motor cortical activity patterns that trigger optogenetic VTA self-stimulation. Learning was accompanied by gradual shaping of these patterns, with participating neurons progressively increasing and aligning their covariance to that of the target pattern. Motor cortex patterns that lead to phasic dopaminergic VTA activity are progressively reinforced and shaped, suggesting a mechanism by which animals select and shape actions to reliably achieve reinforcement.

According to Thorndike's law of effect (*1*), actions that lead to reinforcements are repeated more frequently (*2*). Through repeated attempts, actions are shaped to more directly and reliably achieve reinforcement (*3*, *4*), a process accompanied by the refinement of behavior-specific neural ensembles and activity patterns in motor cortices (*5*–*9*). Learning occurs because neural patterns initiating actions that lead to reinforcement are reentered more often, as supported by neural activity operant conditioning experiments (*10*–*15*).

Reinforcement is thought to rely on the activity of midbrain dopamine neurons. When animals receive reward, dopamine neurons in the ventral tegmental area (VTA) produce a spike burst that encodes the difference between the animal's expected and received rewards (*16*). This reward-prediction error signal is useful for optimizing reward-seeking behavior (*17*, *18*). Indeed, phasic VTA activity constitutes a neural basis of reinforcement, as animals shape their behavior to receive electrical (*19*, *20*) as well as optogenetic (*21*, *22*) VTA self-stimulation.

To test a neural law of effect, we investigated if mice would learn to reenter specific motor cortical patterns to receive dopaminergic VTA self-stimulation (Fig. 1A). We recorded the activity of tens of neurons in primary motor cortex (M1) and used it to trigger optogenetic stimulation of dopaminergic VTA neurons with blue light (*21*). Tyrosine hydroxylase (TH)–Cre mice (*23*) expressing channelrhodopsin-2 (ChR2 group, $n = 10$) in VTA dopaminergic cells were implanted with an optic fiber in the VTA and an electrode array in contralateral M1 layer 5 (Fig. 1B and fig. S1). To control for the effects of viral expression and shining light in the VTA, we expressed yellow fluorescent protein (YFP group, $n = 6$) in Cre-positive mice that underwent the same experimental procedure. Mice were trained to control a brain-machine interface (BMI) that transformed the activity of groups of neurons in M1 into real-time auditory feedback. When mice produced the target neural activity pattern that led to the target tone, they received a train of blue laser pulses, providing phasic stimulation of dopaminergic cells in the VTA. The self-stimulation optogenetic protocol used here has been previously shown to reinforce lever pressing (fig. S2).

This closed-loop self-stimulation paradigm (*24*) provides a principled way to study neural reinforcement, as it assigns chosen recorded neurons ("direct neurons") to drive behavior, defines the transform between neural activity and behavior through the "decoder," and delivers temporally precise reinforcement after target neural activity is produced. Our decoder received input from two arbitrarily selected M1 ensembles of two to four well-isolated single units (see supplementary methods and fig. S3) (*14*, *15*). Two target neural population activity patterns (targets 1 and 2) were specified, which occur with equal frequency in spontaneous activity: Target 1 required the simultaneous positive modulation of ensemble 1 and negative modulation of ensemble 2, whereas target 2 required the reverse modulation (see supplementary methods). The BMI provided optogenetic reinforcement of target 1 only, permitting comparison of the two targets. Further, it provided continuous auditory feedback of neural activity pattern exploration along the task-relevant neural dimension—the differential modulation of ensembles 1 and 2.

We sought to measure how neural reinforcement changes the animals' production of neural activity patterns and resulting occupancy of auditory tones. The initial conditions of learning were established with decoder calibration to set the baseline chance rate of neural activity patterns occupying the tones. During a baseline block preceding each BMI training block, calibration was used to estimate the distribution of ensemble 1 and 2 modulations during spontaneous neural activity while mice freely moved in the behavioral box without receiving auditory feedback or VTA stimulation (Fig. 1C). Each unit's spiking activity was binned in 500-ms bins, and an ensemble's firing-rate modulation was defined as the sum of each unit's median-centered and range-normalized spike count. For each individual ensemble, four modulation states were defined by the 10th, 50th, and 90th percentile of the modulation distribution from the baseline block. The decoder calculated the difference between ensemble 1's and ensemble 2's modulation state for each 500-ms cycle and mapped it to one of seven auditory tones (ranging from 5 to 19 kHz). This daily calibration yielded a Gaussian-like distribution over tones during baseline and ensured that the chance rate of tone occupancy did not change over training days, despite potential day-to-day variability in neural recordings (Fig. 1D). Animals had to produce substantial ensemble modulations to achieve the targets (Fig. 1E). During the BMI training block, neural patterns close to target 1 decreased the tone, whereas neural patterns close to target 2 increased the tone (Fig. 1A). Target achievement resulted in a 1-s playback of the target tone, and only target 1 achievement resulted in phasic VTA stimulation 1.5 s after target hit, consisting of a 14-Hz train delivered for 2 s (Fig. 1F).

We trained animals on four consecutive daily sessions and quantified how reinforcement changed BMI tone distributions relative to session 1 (Fig. 2, A and B). Experimenters were blind to the type of virus injected in the VTA. ChR2 animals changed their target tone occupancy from their baseline bootstrap distribution by sessions 3 and 4, whereas YFP animals showed no preference for target 1 (Fig. 2C). With training, target 1 was occupied significantly more often in ChR2 animals and did not change in YFP animals (Fig. 2D). ChR2 animals increased preference for target 1 versus target 2 (Fig. 2E) and biased their overall distribution toward low-pitch tones close to target 1 and away from high-pitch tones close to target 2 (Fig. 2F). Interestingly, neuroprosthetic-triggered VTA stimulation did not reinforce specific overt movements (*19*, *20*, *22*) or place preference (*21*), suggesting that animals are not simply undergoing motor learning (fig. S4).

Given that the differential modulation between ensembles 1 and 2 shifted toward target 1, we asked more generally how the joint activity of neurons involved in producing the pattern (direct neurons) was shaped by reinforcement. Because the ensembles' simultaneous modulation triggered reinforcement, VTA stimulation might strengthen

[1]Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, Lisbon 1400-038, Portugal. [2]Department of Electrical Engineering and Computer Sciences, University of California–Berkeley, Berkeley, CA 94720, USA. [3]Helen Wills Neuroscience Institute, University of California–Berkeley, Berkeley, CA 94720, USA. [4]Joint Graduate Group in Bioengineering University of California–Berkeley and University of California–San Francisco, Berkeley, CA 94720, USA. [5]Departments of Neuroscience and Neurology, Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY 10032, USA.
*These authors contributed equally to this work. †These authors contributed equally to this work.
‡Corresponding author. Email: rc3031@columbia.edu (R.M.C.);

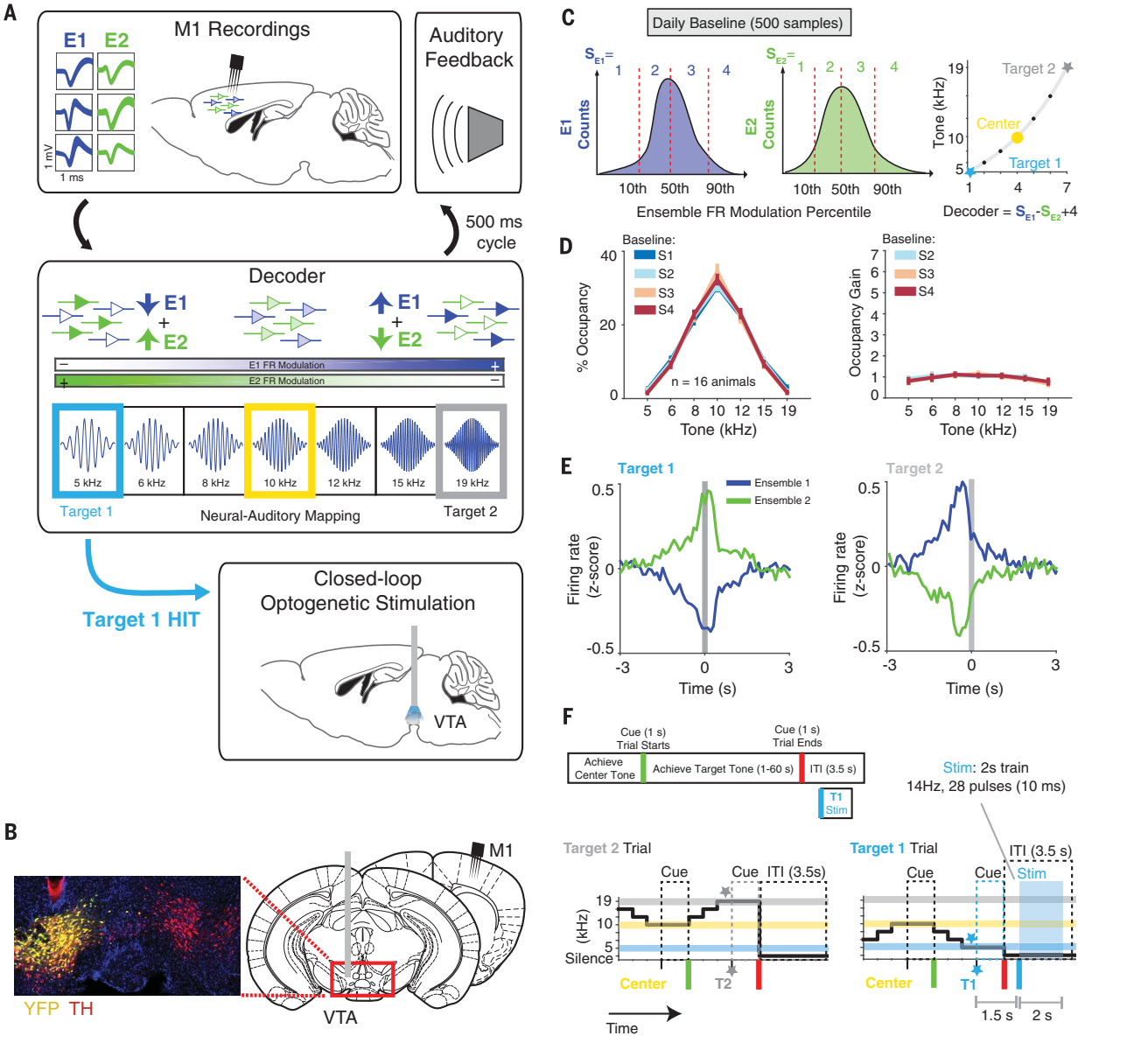


Fig. 1. Closed-loop BMI paradigm for pairing specific motor cortex activity patterns with phasic VTA dopaminergic activity. (**A**) Schematic of the BMI paradigm. Each mouse receives a unilateral microwire array implant in the motor cortex (targeted to layer V) and a contralateral optical fiber implant in the VTA. Recorded single units are arbitrarily assigned into two ensembles (E), and the concomitant increase (up arrow) of one ensemble's activity and decrease (down arrow) in the other ensemble's activity drives the decoder to change the auditory tone produced every 500 ms. The rare, lowest-pitch tone triggers phasic optical stimulation to the VTA, whereas the rare, highest-pitch tone serves as a control. Solid triangles indicate neurons with positively modulated firing rate; open triangles indicate neurons with negatively modulated firing rate. Yellow color indicates the center-pitch tone. FR, firing rate. (**B**) Coronal brain slice depicting viral infection specific to the dopaminergic cells of the VTA. The immunohistochemistry labels for tyrosine hydroxylase (TH, red) and the Cre-dependent fluorescent protein (YFP, yellow) are shown. (**C**) BMI decoder calibration. For every session (S) during the baseline period, 500 samples of 500-ms spike counts are collected from spontaneous neural activity as the mouse freely behaves in the box with no task or auditory tones. Each ensemble's firing rate modulation is defined as the sum of the member neurons' normalized spike counts (mean-centered, range-normalized) and then quantized into four activation states. The decoder's state is the difference between ensemble 1's and ensemble 2's activation state and is mapped into one of seven tones. The stars indicate target tones. (**D**) BMI calibration on baseline period spontaneous neural activity results in a Gaussian-like distribution over tones, such that target 1 (5 kHz) and target 2 (19 kHz) are rare. The mean and SEM baseline distribution for each session is plotted on the left, averaged over all animals. Baseline distributions show no change from session 1, as shown on the right. (**E**) Ensemble 1 and 2 firing rate modulation before target 1 and target 2 hits, averaged over all recorded cells and sessions. (**F**) Task schematic. Trial structure is the same for target 1 and target 2, except that a target 1 hit results in phasic VTA stimulation (2-s train of 14 Hz pulses with 10-ms width). ITI, intertrial interval.

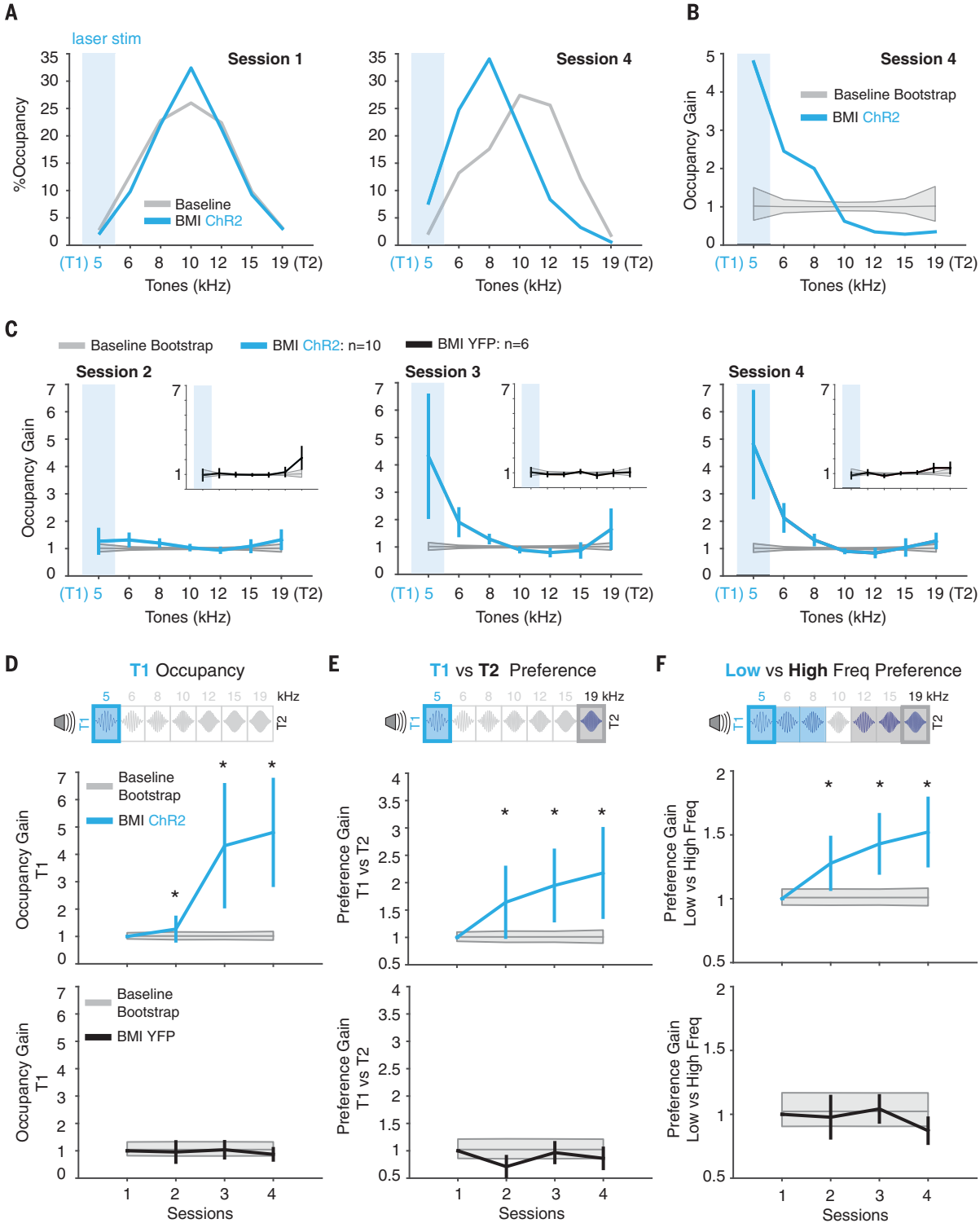


Fig. 2. Target pattern reentrance increases during VTA optogenetic self-stimulation. (**A**) Distribution of the percent of time that each tone was occupied during baseline (gray) and BMI (cyan) blocks of session 1 (left) and session 4 (right) in one mouse. (No tones were actually played during the baseline block.) T1, target 1; T2, target 2. (**B**) Quantification of the behavioral changes between sessions 1 and 4. The session 4 occupancy gain (cyan) is the session 4 BMI distribution normalized to the session 4 baseline distribution, then normalized to the session 1 ratio. For (B) to (F), the 95% confidence interval for the baseline bootstrap distribution is plotted in gray (see supplementary methods). To generate the bootstrap distribution, the BMI session was simulated 10,000 times as though neural activity were drawn from that session's baseline period. (**C**) The occupancy gain over sessions 2 through 4. For (C) to (F), mean and SEM over ChR2 animals ($n = 10$) are shown in cyan and over YFP animals ($n = 6$) are shown in black. By session 4, the behavioral changes were statistically different across tones for ChR2 but not YFP [repeated measures analysis of variance (ANOVA): ChR2, $F_{6,48} = 3.46$, $P = 6.4 \times 10^{-3}$; YFP, $F_{6,30} = 0.96$, $P = 0.47$]. In session 4, 5 kHz (target 1) was significantly different from all tones from 8 to 19 kHz (Tukey's post hoc multiple comparisons test). (**D**) Top: The occupancy gain for 5 kHz (target 1) over sessions is shown. Middle: ChR2 (cyan) were significantly larger than bootstrap from sessions 2 through 4 (session 2, $P = 1.2 \times 10^{-3}$; session 3, $P < 1 \times 10^{-5}$; session 4, $P < 1 \times 10^{-5}$). Bottom: YFP (black) were never significantly larger than bootstrap. (**E**) Top: The preference gain for 5 kHz (target 1) versus 19 kHz (target 2) is plotted over sessions. Middle: ChR2 (cyan) were significantly larger than bootstrap after session 1 ($P < 1 \times 10^{-5}$ for sessions 2 through 4). Bottom: YFP (black) were never significantly larger than bootstrap. (**F**) Top: The preference gain for low-pitch tones (5 to 8 kHz, close to target 1) versus high-pitch tones (12 to 19 kHz, close to target 2) over sessions is shown. Middle: ChR2 (cyan) were significantly larger than bootstrap after session 1 ($P < 1 \times 10^{-5}$ for sessions 2 through 4). Bottom: YFP (black) were never significantly larger than bootstrap. For (D) to (F), an asterisk indicates that the population average is significantly larger than the baseline bootstrap distribution.

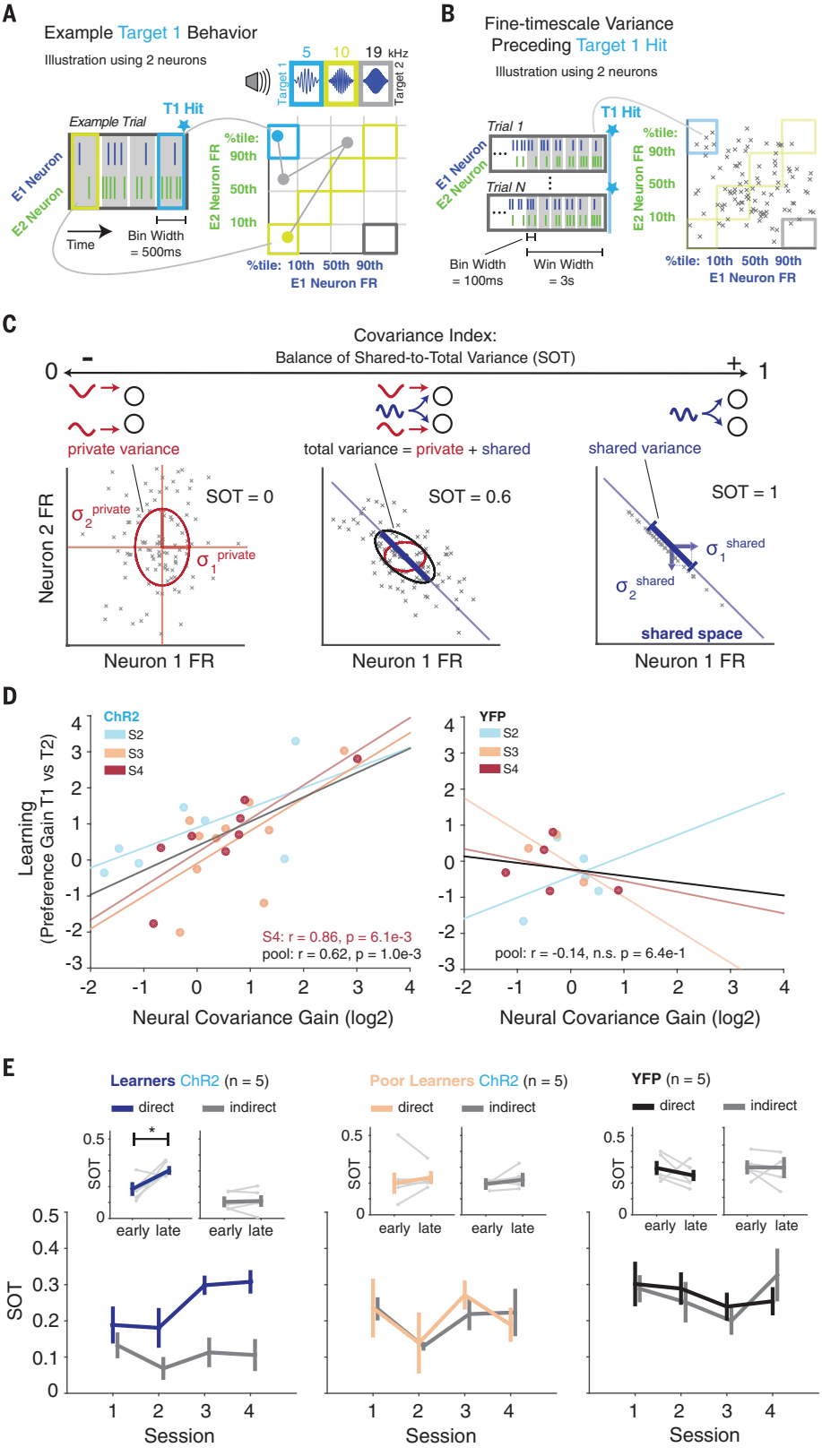


Fig. 3. Learning correlates with an increase in covariance of the neurons that produce the target pattern. (**A**) The decoder maps spike counts in 500-ms bins into quantizations of (ensemble 1, ensemble 2) space. Neural activity can take multiple routes to achieve target 1. (**B**) Analysis of variance of spike counts with 100-ms bins in a 3-s window preceding target hit. "x" indicates a spike count vector at one time point. (**C**) Factor analysis was used to analyze the ratio of shared variance to total variance (SOT), which ranges from 0 to 1, for the full population controlling the BMI. A two-neuron illustration shows a neural solution with SOT = 0, 0.6, and 1. (**D**) Correlation of change in shared variance before target 1 hit (neural covariance gain) with change in preference for target 1 over target 2 (learning), over sessions 2, 3, and 4. ChR2 animals (left) showed a significant correlation [ChR2 S4: correlation coefficient (r) = 0.86, $P = 6.1 \times 10^{-3}$; ChR2 pool S3, S4: $r = 0.71$, $P = 1.0 \times 10^{-3}$; ChR2 pool S2, S3, S4: $r = 0.62$, $P = 9.8 \times 10^{-4}$; ChR2 S3: $r = 0.60$, $P = 6.5 \times 10^{-2}$; ChR2 S2: $r = 0.62$, $P = 1.3 \times 10^{-1}$], whereas YFP animals (right) showed no correlation (YFP pool S2, S3, S4: $r = -0.14$, n.s. $P = 6.4 \times 10^{-1}$; YFP S4: $r = -0.32$, $P = 6.0 \times 10^{-1}$; YFP S3: $r = -0.69$, $P = 5.1 \times 10^{-1}$; YFP S2: $r = 0.37$, $P = 5.4 \times 10^{-1}$). n.s., not significant. (**E**) SOT of direct and indirect neurons over sessions for ChR2 learners (left, $n = 5$), ChR2 poor learners (middle, $n = 5$), and YFP subjects (right, $n = 5$). ChR2 learners individually showed significant preference gain for target 1 versus target 2 in both sessions 3 and 4. ChR2 poor learners constitute the remaining animals who as a population showed significant target 1 occupancy gain on sessions 3 and 4. For direct neurons, ChR2 animals' and ChR2 learners' SOT increased from early (sessions 1 and 2 pooled) to late training (sessions 3 and 4 pooled), whereas ChR2 poor learners and YFP did not (one-sided rank sum test; ChR2, early < late, $P = 1.7 \times 10^{-2}$; ChR2 learners, early < late, $P = 1.6 \times 10^{-2}$; ChR2 poor learners, early < late, n.s. $P = 2.1 \times 10^{-1}$; YFP, early < late, n.s. $P = 8.3 \times 10^{-1}$). For indirect neurons, SOT showed no change for all groups (ChR2 learners: early < late, n.s. $P = 4.3 \times 10^{-1}$; ChR poor learners, early < late, n.s. $P = 2.7 \times 10^{-1}$; YFP, early < late, n.s. $P = 7.1 \times 10^{-1}$). Traces in the insets show the average of each animal's SOT in sessions 1 and 2 (early) versus the average of sessions 3 and 4 (late). Error bars indicate mean ± SEM. The asterisk indicates that the population average is significantly larger than the baseline bootstrap distribution.

shared inputs to direct neurons and thus increase covariance over learning (*13*). We used factor analysis (FA) to partition fine–time scale neural variance arising from two sources: private inputs to each cell, which drive independent firing (private variance), and shared inputs, which drive multiple cells simultaneously (shared variance). Neural variance changes were not demanded by our task, as subjects could use neural activity drawn from any distribution to ultimately hit target 1 (Fig. 3A). We analyzed fine–time scale spike counts (100-ms bins) in a 3-s window preceding target hit (Fig. 3B). FA models population spike counts $x = \mu + x^{private} + x^{shared}$ as the sum of a mean firing rate μ; private variation

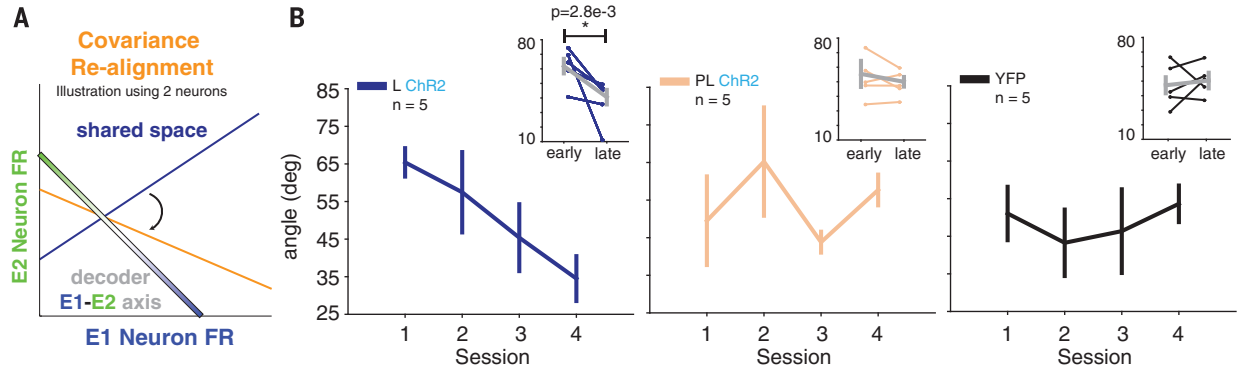


Fig. 4. Covariance of the neurons that produce the target pattern gradually aligns to the decoder. (**A**) Analysis of shared variance alignment with the decoder's ensemble 1 and ensemble 2 assignments by using the angle between the shared space and the decoder's "ensemble 1 minus ensemble 2" axis. Curved arrow indicates rotation of the shared space to align with the decoder. (**B**) The angle between shared variance and the decoder axis decreased for ChR2 learners (left) but not for poor learners (middle) and YFP (right) (one-sided rank sum test comparing sessions 1 and 2 to sessions 3 and 4; ChR2 learner, late < early, $P = 2.8 \times 10^{-3}$; ChR2 poor learner, late < early, n.s. $P = 3.7 \times 10^{-1}$; YFP, late < early, n.s. $P = 7.5 \times 10^{-1}$). Traces in the insets show the average of each animal's angle in sessions 1 and 2 (early) versus the average of sessions 3 and 4 (late). Error bars indicate mean ± SEM. The asterisk indicates that the population average is significantly larger than the baseline bootstrap distribution.

$x^{private}$, which is uncorrelated across neurons; and shared variation $x^{shared} = Uz$, which is driven by latent shared inputs z through the linear factors U. Because $x^{private}$ and x^{shared} are independent, the total covariance matrix $\Sigma^{total} = \Sigma^{private} + \Sigma^{shared}$ is decomposed into the sum of a diagonal private covariance matrix $\Sigma^{private}$ and a low-rank shared covariance matrix Σ^{shared}. Geometrically, private variance spans all of the high-dimensional population activity space for which each neuron's activity is one dimension, whereas shared variance is constrained to a low-dimensional "shared space" because there are fewer shared inputs than neurons. The number of shared dimensions was fit by using standard model selection (fig. S5) by maximizing cross-validated log likelihood (*13*, *25*–*28*).

We assessed neural coordination with a covariance index defined as the ratio of the shared variance to total variance averaged over neurons (SOT) (Fig. 3C). Although Fig. 3, A to C, uses two neurons for illustration, we emphasize that FA was applied to the joint activity of all neurons used to control the BMI (ranging from four to eight). We then asked if learning, defined as the proportion of hits of target 1 versus target 2 normalized to session 1, was correlated with the increase in covariance, defined as the SOT normalized to session 1. The increase in covariance correlated with learning in ChR2, but not YFP, animals (Fig. 3D). This correlation became stronger as learning progressed.

These data suggest that the degree of learning related to the degree of neural variance changes. To further dissect this, we analyzed ChR2 animals and found two groups distinguished by their degree of learning (fig. S6). Each individual of the learner group ($n = 5$) showed statistically significant preference for target 1 versus target 2 for both sessions 3 and 4. The poor learner group ($n = 5$), as a population, showed an increase in target 1 occupancy but did not improve preference for target 1 over target 2 (fig. S6). Learners significantly increased their covariance index over training, whereas poor learners and YFP did not (Fig. 3E and figs. S7 and S8A). This effect was ensemble specific, as only neurons controlling the BMI (direct neurons) increased their covariance index, whereas other recorded neurons (indirect neurons) did not (Fig. 3E and figs. S9 and S10).

Finally, we asked whether dopaminergic self-stimulation shaped the neural covariance to more easily achieve the target pattern. Only neural variance that causes differential modulation between ensembles 1 and 2 can change the feedback tone and contribute to target achievement, corresponding to variance that is aligned with the decoder's "ensemble 1 minus ensemble 2" axis (Fig. 4A). We analyzed the relationship between shared neural variance and the decoder by calculating the angle between the shared space and the decoder axis. The angle between the shared space and the decoder axis decreased significantly for learners but not for poor learners and YFP (Fig. 4B and fig. S8B).

The results presented here show that mice reenter specific neural patterns that trigger dopaminergic VTA self-stimulation more often as training progresses. Dopaminergic self-stimulation not only increases the reentry of a target pattern, which may have been strongly predicted on the basis of previous studies, but further shapes the distribution of activity patterns to more directly achieve the target pattern. The covariance increased specifically between direct neurons and gradually became aligned with the decoder. Individual neuron firing properties did not correlate with learning (fig. S11), highlighting that it was the specific pattern that was reinforced. This reinforcement of specific neural ensembles and activity patterns extends recent work showing individual neuron conditioning through electrical self-stimulation of the nucleus accumbens (*29*). Although it may be difficult to completely rule out that very subtle movements that lead to the desired patterns of activity are being reinforced, we showed that, in this paradigm, there is no reinforcement of overt movements over BMI learning (fig. S4). Still, these results may have implications for motor reinforcement, in which actions are selected more often and optimized over iterations to more directly achieve reinforcements.

In these experiments, subjects learned to produce neural patterns de novo, which leverages different mechanisms from BMI learning experiments in which subjects adapted to decoder perturbations. BMI-experienced subjects learn to control a decoder by selecting activity patterns from their existing shared space (*28*). By contrast, our learners initially exhibit little shared variance, and this shared variance is misaligned with the decoder. Thus, they likely select patterns from their high-dimensional private variance, gradually developing and realigning shared variance with learning (*13*). Analysis and modeling indicate that private variance is useful for broad exploration of population activity space (*13*) and for learning each neuron's contributions to achieving a goal (*30*, *31*), possibly permitting the selective increase of direct neurons' covariation index over indirect neurons. The difference between learners and poor learners could depend on the probability of the direct neurons receiving common anatomical input, or on the plasticity of common inputs to the direct neurons.

It is unlikely that VTA stimulation directly modulated activity and plasticity in M1 through monosynaptic projections because we stimulated the VTA contralateral to our M1 recordings, and most projections are unilateral. Indeed, VTA stimulation did not induce any observable changes in the mean firing rates of M1 neurons (fig. S12). Thus, M1 reinforcement is likely driven by inputs from and plasticity in broader networks, such as cortico-basal ganglia circuits. Cortico-striatal plasticity is modulated by dopamine (*32*, *33*) and is

necessary for motor and neuroprosthetic learning (*5*, *14*, *34*). Actor-critic reinforcement learning models (*35*, *36*) suggest two sites for VTA-modulated plasticity: the dorsal striatum (actor), which contributes to selection of actions (M1 neural activity patterns), and the ventral striatum (critic), which may evaluate actions on the basis of the value of the environmental states reached (auditory feedback). Plasticity in dorsal striatum could be mediated by glutamatergic input from contralateral M1 and dopaminergic input signaling reward from the VTA (*37*), enabling adaptation of the policy for reentering M1 activity patterns. Plasticity in ventral striatum (*32*) could be mediated by strong bidirectional connections with the VTA, enabling adaptation of the auditory tones' value.

In addition, VTA stimulation may have indirectly guided motor cortical plasticity. As animals acquire motor skills and consolidate cortical activity patterns, motor memories are encoded in the formation of lasting dendritic spine ensembles (*8*, *38*–*40*). Further, reinforcement guides the reactivation of neurons during sleep (*41*), leading to the formation of dendritic spines (*42*) as well as the identification of neurons responsible for achieving a target pattern (*41*). Thus, our observed changes in shared variance could also reflect sleep-dependent changes in motor cortical synaptic connectivity. Recent modeling work shows that excitation-inhibition–balanced spiking networks with clustered connectivity exhibited prominent low-dimensional shared variance, whereas nonclustered networks exhibited weak, high-dimensional shared variance (*43*).

Our results provide causal evidence for a neural law of effect, describing how the brain selects and shapes neural activity patterns through neural reinforcement. As Skinner noted, selection by consequence is a mechanism driving the evolution of living things, from species to societies to behavior (*44*). Our results help uncover how selection by consequence operates on neural activity in the brain (*45*).

REFERENCES AND NOTES

1. E. L. Thorndike, *Psychol. Rev.* **2**, 1–107 (1898).
2. B. F. Skinner, *The Behavior of Organisms: An Experimental Analysis* (Appleton-Century, Oxford, 1938).
3. R. G. Cohen, D. Sternad, *Exp. Brain Res.* **193**, 69–83 (2009).
4. L. Shmuelof, J. W. Krakauer, P. Mazzoni, *J. Neurophysiol.* **108**, 578–594 (2012).
5. R. M. Costa, D. Cohen, M. A. L. Nicolelis, *Curr. Biol.* **14**, 1124–1134 (2004).
6. T. D. Barnes, Y. Kubota, D. Hu, D. Z. Jin, A. M. Graybiel, *Nature* **437**, 1158–1161 (2005).
7. V. Y. Cao *et al.*, *Neuron* **86**, 1385–1392 (2015).
8. A. J. Peters, S. X. Chen, T. Komiyama, *Nature* **510**, 263–267 (2014).
9. B. P. Ölveczky, T. M. Otchy, J. H. Goldberg, D. Aronov, M. S. Fee, *J. Neurophysiol.* **106**, 386–397 (2011).
10. E. E. Fetz, *Science* **163**, 955–958 (1969).
11. J. M. Carmena *et al.*, *PLOS Biol.* **1**, E42 (2003).
12. K. Ganguly, J. M. Carmena, *PLOS Biol.* **7**, e1000153 (2009).
13. V. R. Athalye, K. Ganguly, R. M. Costa, J. M. Carmena, *Neuron* **93**, 955–970.e5 (2017).
14. A. C. Koralek, X. Jin, J. D. Long II, R. M. Costa, J. M. Carmena, *Nature* **483**, 331–335 (2012).
15. K. B. Clancy, A. C. Koralek, R. M. Costa, D. E. Feldman, J. M. Carmena, *Nat. Neurosci.* **17**, 807–809 (2014).
16. W. Schultz, P. Dayan, P. R. Montague, *Science* **275**, 1593–1599 (1997).
17. R. S. Sutton, A. G. Barto, *Reinforcement Learning: An Introduction* (MIT Press, 1998).
18. E. E. Steinberg *et al.*, *Nat. Neurosci.* **16**, 966–973 (2013).
19. J. Olds, P. Milner, *J. Comp. Physiol. Psychol.* **47**, 419–427 (1954).
20. D. Corbett, R. A. Wise, *Brain Res.* **185**, 1–15 (1980).
21. H.-C. Tsai *et al.*, *Science* **324**, 1080–1084 (2009).
22. I. B. Witten *et al.*, *Neuron* **72**, 721–733 (2011).
23. S. Gong *et al.*, *J. Neurosci.* **27**, 9817–9823 (2007).
24. L. Grosenick, J. H. Marshel, K. Deisseroth, *Neuron* **86**, 106–139 (2015).
25. A. P. Dempster, N. M. Laird, D. B. Rubin, *J. R. Stat. Soc. B* **39**, 1–38 (1977).
26. B. S. Everitt, *An Introduction to Latent Variable Models* (Chapman and Hall, London, 1984).
27. B. M. Yu *et al.*, *J. Neurophysiol.* **102**, 614–635 (2009).
28. P. T. Sadtler *et al.*, *Nature* **512**, 423–426 (2014).
29. R. W. Eaton, T. Libey, E. E. Fetz, *J. Neurophysiol.* **117**, 1112–1125 (2017).
30. R. Legenstein, S. M. Chase, A. B. Schwartz, W. Maass, *J. Neurosci.* **30**, 8400–8410 (2010).
31. R. Héliot, K. Ganguly, J. Jimenez, J. M. Carmena, *IEEE Trans. Syst. Man Cybern. B Cybern.* **40**, 1387–1397 (2010).
32. S. Yagishita *et al.*, *Science* **345**, 1616–1620 (2014).
33. W. Shen, M. Flajolet, P. Greengard, D. J. Surmeier, *Science* **321**, 848–851 (2008).
34. F. J. Santos, R. F. Oliveira, X. Jin, R. M. Costa, *eLife* **4**, e09423 (2015).
35. Y. Takahashi, G. Schoenbaum, Y. Niv, *Front. Neurosci.* **2**, 86–99 (2008).
36. J. O'Doherty *et al.*, *Science* **304**, 452–454 (2004).
37. M. W. Howe, D. A. Dombeck, *Nature* **535**, 505–510 (2016).
38. M. Fu, X. Yu, J. Lu, Y. Zuo, *Nature* **483**, 92–95 (2012).
39. T. Xu *et al.*, *Nature* **462**, 915–919 (2009).
40. A. Hayashi-Takagi *et al.*, *Nature* **525**, 333–338 (2015).
41. T. Gulati, L. Guo, D. S. Ramanathan, A. Bodepudi, K. Ganguly, *Nat. Neurosci.* **20**, 1277–1284 (2017).
42. G. Yang *et al.*, *Science* **344**, 1173–1178 (2014).
43. R. C. Williamson *et al.*, *PLOS Comput. Biol.* **12**, e1005141 (2016).
44. B. F. Skinner, *Science* **213**, 501–504 (1981).
45. R. M. Costa, *Curr. Opin. Neurobiol.* **21**, 579–586 (2011).

ACKNOWLEDGMENTS

We thank A. Castro for the motor reinforcement experiments; A. Koralek, P. Khanna, and R. Neely for helpful discussions; and A. Vaz for animal colony management. This work was supported by a NSF Graduate Research Fellowship to V.R.A.; grants from the NSF (CBET-0954243 and EFRI-M3C 1137267) and Office of Naval Research (N00014-15-1-2312) to J.M.C.; and grants from the European Research Area ERA-NET, European Research Council (COG 617142), and Howard Hughes Medical Institute (IEC 55007415) to R.M.C. Data presented in this paper can be found on figshare at https://doi.org/10.6084/m9.figshare.5687101.v1.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1024/suppl/DC1
Materials and Methods
Figs. S1 to S12
References (*46*–*49*)

8 August 2017; accepted 8 January 2018
10.1126/science.aao6058

EVOLUTION

Incomplete host immunity favors the evolution of virulence in an emergent pathogen

Arietta E. Fleming-Davies,[1,2,3]* Paul D. Williams,[4]*† André A. Dhondt,[5] Andrew P. Dobson,[4,6] Wesley M. Hochachka,[5] Ariel E. Leon,[2] David H. Ley,[7] Erik E. Osnas,[8] Dana M. Hawley[2]‡

Immune memory evolved to protect hosts from reinfection, but incomplete responses that allow future reinfection may inadvertently select for more-harmful pathogens. We present empirical and modeling evidence that incomplete immunity promotes the evolution of higher virulence in a natural host-pathogen system. We performed sequential infections of house finches with *Mycoplasma gallisepticum* strains of various levels of virulence. Virulent bacterial strains generated stronger host protection against reinfection than less virulent strains and thus excluded less virulent strains from infecting previously exposed hosts. In a two-strain model, the resulting fitness advantage selected for an almost twofold increase in pathogen virulence. Thus, the same immune systems that protect hosts from infection can concomitantly drive the evolution of more-harmful pathogens in nature.

Imperfect vaccines promote the evolution of more-harmful pathogens by creating a fitness advantage for virulent strains in vaccinated hosts, such as high rates of infectivity or transmission (*1*–*5*). By preventing disease-induced host death and subsequent removal of virulent strains from a population, imperfect vaccines also reduce the overall costs of virulence to pathogens (*5*). We asked whether incomplete immune responses to natural infections can similarly favor the evolution of more-virulent pathogen strains that cause greater host mortality. The bacterial pathogen *Mycoplasma gallisepticum* emerged in the 1990s in free-living house finches (*Haemorhous mexicanus*), causing severe conjunctivitis that indirectly reduces finch survival via visual impairment and reduced ability to escape predators (*6*, *7*). After emergence in the eastern United States, *M. gallisepticum* spread throughout the house finch range (*8*), transmitted by direct contact and contaminated surfaces such as bird feeders (*9*). Soon after the pathogen became endemic on each coast of the United States, the virulence of circulating *M. gallisepticum* strains rapidly increased, as measured by disease severity produced in immunologically naïve hosts (*10*). More than half of free-living house finches can recover from *M. gallisepticum* infection (*6*), generating pools of recovered hosts in wild populations. Furthermore, recovered hosts show strong but incomplete immune protection and thus can become reinfected with homologous or heterologous pathogen strains (*11*–*13*). In our study, we tested whether incomplete host immunity drives the evolution of greater virulence in this system.

We used sequential-inoculation experiments to quantify how incomplete immunity generated from experimental prior exposure alters host responses relevant for pathogen fitness. House finches naïve to *M. gallisepticum* at capture (n = 120 finches) were individually housed and sequentially exposed to pairs of *M. gallisepticum* strains that either were identical (homologous) or had different levels of virulence (heterologous), with clinical recovery between exposures. We performed two identical experiments in successive years, each using three distinct pathogen strains in a completely randomized design (Tables 1 and 2). Positive controls received sterile medium during primary inoculation and thus had no pathogen exposure before secondary inoculation with one of the six strains. We quantified strain virulence as the degree of within-host replication (ε in Tables 1 and 2) using conjunctival pathogen loads measured by quantitative polymerase chain reaction (qPCR) (*10*) across 8 weeks after primary inoculation of immunologically naïve birds (see supplementary materials). To examine how the virulence of the primary and secondary strains affected host responses relevant for pathogen fitness, we measured conjunctivitis severity and pathogen load for 5 weeks after secondary inoculation. Conjunctivitis severity, which correlates with disease-induced mortality risk in the wild (*6*, *7*), was scored on an ordinal scale from 0 to 3, and conjunctival pathogen load was quantified by qPCR (*10*).

First, we found that hosts with prior experimental exposure to any *M. gallisepticum* strain showed reduced severity of the clinical signs that predict mortality risk in the wild (*6*, *7*) compared with hosts with no prior exposure (Fig. 1A). Thus, consistent with findings from earlier work using vaccines (*5*), host immunity reduced the costs of virulence to the pathogen by protecting hosts

Table 1. Relative virulence and $\log_{10}$ conjunctival pathogen loads (means ± SEM) for experiment one. All birds received primary and secondary inoculations with either sterile medium or one of three *M. gallisepticum* strains that varied in virulence (*10*). Strains are ordered from low to high virulence (left to right and top to bottom), and relative virulence terms describe the primary exposure strain relative to the secondary exposure strain (e.g., "less virulent" indicates primary exposure to a less virulent strain than the secondary strain). Numbers in column headings indicate strain virulence estimated from a separate data set for naïve hosts (see supplementary materials). Strain virulence values (ε) are linear model coefficients (for a linear mixed-effects model fitting strain effects in naïve birds using data from all experimental strains; likelihood ratio = 172.56, df = 6, 124, $P < 0.001$), whereas values in the table are raw means of pathogen load (n = 4 to 5 birds for each group). Pathogen loads tend to increase top to bottom (with increasing virulence of the secondary strain) but decrease left to right within a row (with increasing virulence of the primary strain). N/A, not applicable.

	Relative virulence and mean pathogen load for primary exposure			
Secondary exposure	**Sterile medium (control)**	**CA2006 low (ε = 1.41 ± 0.34)**	**VA1994 intermediate (ε = 3.31 ± 0.34)**	**NC2006 high (ε = 4.72 ± 0.34)**
CA2006	No prior exposure 1.54 ± 0.23	Homologous 0.47 ± 0.17	More virulent 0.20 ± 0.13	More virulent 0.27 ± 0.14
VA1994	No prior exposure 3.29 ± 0.35	Less virulent 1.11 ± 0.31	Homologous 0.62 ± 0.20	More virulent 1.10 ± 0.18
NC2006	No prior exposure 4.71 ± 0.52	Less virulent 3.30 ± 0.39	Less virulent 1.78 ± 0.18	Homologous 1.02 ± 0.17
Sterile medium (control)	Negative control 0.044 ± 0.027	N/A	N/A	N/A

[1]Department of Biology, University of San Diego, San Diego, CA 92110, USA. [2]Department of Biological Sciences, Virginia Tech, Blacksburg, VA 24061, USA. [3]Department of Biology, Radford University, Radford, VA 24141, USA. [4]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA. [5]Lab of Ornithology, Cornell University, Ithaca, NY 14850, USA. [6]Santa Fe Institute, Santa Fe, NM 87501, USA. [7]Department of Population Health and Pathobiology, College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27607, USA. [8]U.S. Fish and Wildlife Service, Anchorage, AK 99503, USA.
*These authors contributed equally to this work. †Deceased.
‡Corresponding author. Email: hawleyd@vt.edu

from the disease-induced mortality that prevents onward pathogen transmission and thus reduces pathogen fitness to zero. Second, we observed the greatest reduction in pathogen load and clinical signs and, thus, the strongest protection against reinfection in hosts previously exposed to higher-virulence strains (Fig. 1 and Tables 1 and 2). Primary exposure to a homologous strain also generated strong host protection (Fig. 1), suggesting that adaptive immune responses, either alone or in combination with innate priming mechanisms (*14*), likely underlie the detected incomplete protection against reinfection (see also fig. S1).

We used the empirical responses to secondary inoculation to fit two key pathogen-associated traits, disease-induced mortality and susceptibility, as continuous functions of the virulence of both the primary and secondary strains (see supplementary materials). Disease-induced mortality rates were inferred to scale linearly with conjunctivitis severity, which predicts mortality risk in the wild (*7*); susceptibility to infection was inferred from the presence of a pathogen load above 10^2 copies in the conjunctiva at any post-inoculation qPCR sampling point. As expected, disease-induced mortality and susceptibility increased with the virulence of the currently infecting strain (Fig. 2 and tables S1 and S2). Thus, virulent strains are better able to successfully infect hosts but also cause higher disease-induced mortality than less virulent strains. However, the host immune status generated by primary treatment strongly modifies the degree of disease-induced mortality and susceptibility for all strains. Hosts with prior exposure to high-virulence strains (Fig. 2, red lines) had the lowest disease-induced mortality and susceptibility among all primary treatment groups. Thus, by generating stronger host protection during primary infection (Fig. 1), high-virulence strains effectively exclude low-virulence strains from future infections of that host.

We next asked how the observed relationship between virulence and protection against reinfection alters the optimal or evolutionarily stable level of virulence in a pathogen population (Fig. 3). We used a two-strain SIRS (susceptible-infected-recovered-susceptible) model with the

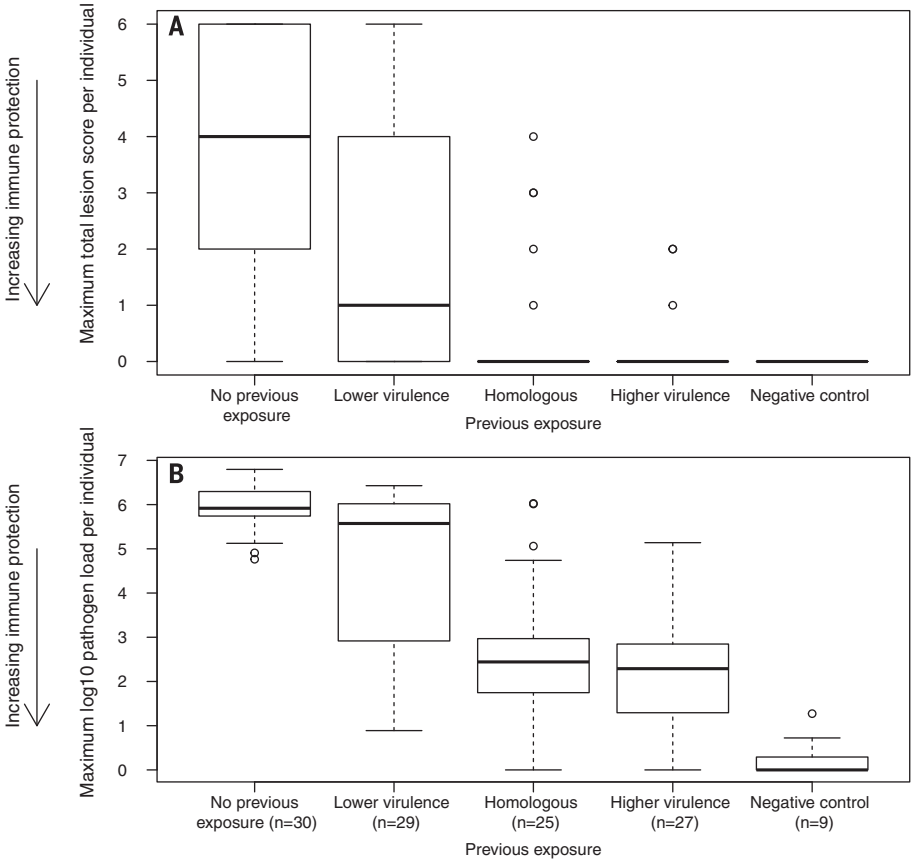


Fig. 1. Lesion scores and pathogen loads for different exposure treatments. Clinical signs of conjunctivitis (**A**), which predict mortality in the wild, and conjunctival pathogen loads (**B**) in hosts with no previous exposure (positive controls) or previous exposure to *M. gallisepticum* strains that either were homologous or had different levels of virulence. Negative controls received sterile medium for both exposures. Host protection was strongest when the primary exposure strain was of higher virulence than the secondary strain, indicating an effect of strain virulence separate from strain homology effects. Box plots show the maximum observed lesion score (on a scale of 0 to 3 per eye, summed within sampling data for a maximum score of 6) or the conjunctival pathogen load for each of 120 individuals from six postexposure measurements. The levels of virulence of the previous-exposure strains are grouped categorically here for clarity (differences among treatment groups were analyzed by the Kruskal-Wallis test; eye score, $H = 68.16$, df = 4, $P < 0.0001$; pathogen load, $H = 76.80$, df = 4, $P < 0.0001$), but virulence was treated as a continuous variable in the model and analysis. *n*, number of individuals.

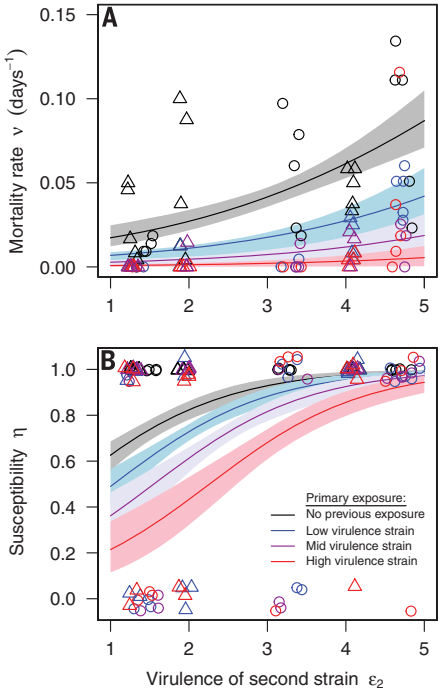


Fig. 2. Model parameters as functions of pathogen virulence, fit to empirical data. Higher virulence of the currently infecting (i.e., secondary) strain is associated with higher host mortality (**A**) and host susceptibility (**B**), but host prior exposure (colored lines) reduces disease-associated mortality and susceptibility. Prior exposure to a high-virulence strain provides the most protection and thus results in the lowest host mortality and susceptibility. The effect of virulence of the primary strain was fit as a continuous function, but to visualize effects, lines and points show three categories of primary strain virulence: low (blue), intermediate (purple), or high (red). Shading around lines denotes bootstrapped 95% confidence intervals incorporating error in virulence estimates. Circles and triangles represent data from experiments one and two, respectively. Mortality rates (A) were inferred to scale linearly with conjunctival lesion scores; susceptibility (B) was fit to binomial infection status (yes or no) inferred from conjunctival pathogen presence. Because distinct strains were used in each experiment (Tables 1 and 2), lines show the function averaged for the two experimental strains in each virulence category (e.g., low virulence). See fig. S4 for functions and data separated by experiment.

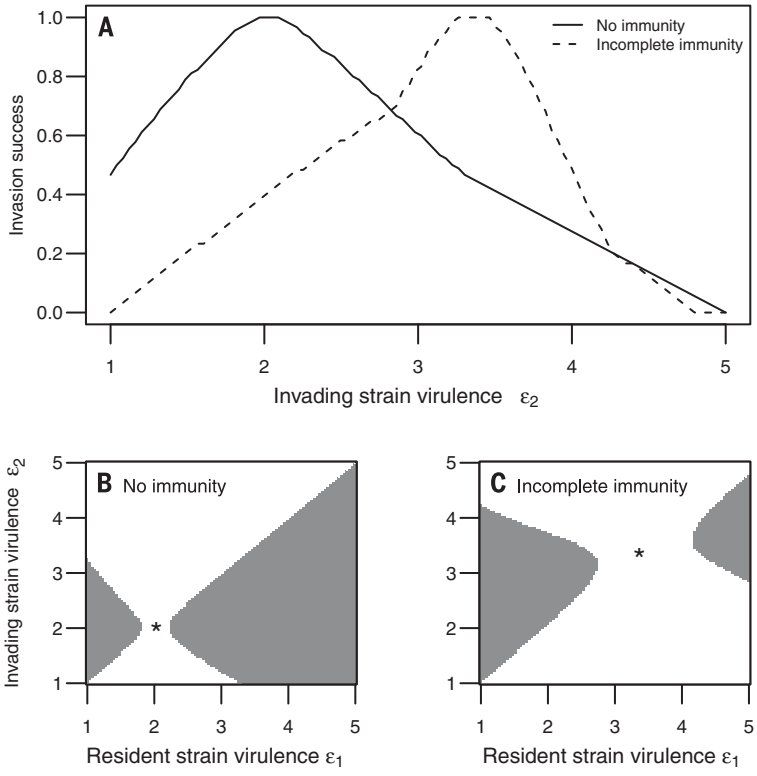


Fig. 3. Optimal virulence in models with no immunity and incomplete immunity. (**A**) The empirically observed effects of incomplete immunity result in an almost twofold increase in optimal virulence relative to that in a model with no immunity. Pathogen fitness is measured as invasion success, computed from the proportion of resident pathogen parameter space over which an invader of a given virulence level was able to successfully displace the resident pathogen in the pairwise invasibility plots (PIPs) for each model (**B** and **C**), scaled to a maximum invasion value of 1 [see supplementary materials and (*35*)]. Shaded areas in the PIPs show parameter space for which a new mutant introduced at very low densities was able to invade a population with equilibrium densities of the resident strain. Nonshaded areas indicate parameter space where the resident strain could not be competitively displaced, and the asterisks in the center of this space thus mark the evolutionarily stable strategy for each model. Susceptibility (η) and mortality (υ) were both continuous functions of strain virulence (ε) (Fig. 2); all other parameters were held constant (equations S2).

Table 2. Relative virulence and log_{10} conjunctival pathogen loads (means ± SEM) for experiment two. The experimental design for experiment two was identical to that for experiment one (Table 1) but used distinct strains. Strains are ordered from low to high virulence (left to right and top to bottom), and relative virulence terms describe the primary exposure strain relative to the secondary exposure strain. As in Table 1, pathogen loads tend to increase top to bottom (with increasing virulence of the secondary strain) but decrease left to right within a row (with increasing virulence of the primary strain). N/A, not applicable.

	Relative virulence and mean pathogen load for primary exposure			
Secondary exposure	**Sterile medium (control)**	**CA2009 low (ε = 1.28 ± 0.34)**	**NC1995 intermediate (ε = 1.95 ± 0.35)**	**VA2013 high (ε = 4.08 ± 0.34)**
CA2009	No prior exposure 2.82 ± 0.33	Homologous 0.68 ± 0.16	More virulent 1.14 ± 0.23	More virulent 1.05 ± 0.20
NC1995	No prior exposure 3.67 ± 0.91	Less virulent 0.86 ± 0.33	Homologous 0.56 ± 0.08	More virulent 0.74 ± 0.14
VA2013	No prior exposure 3.75 ± 0.37	Less virulent 3.00 ± 0.11	Less virulent 2.64 ± 0.23	Homologous 1.51 ± 0.47
Sterile medium (control)	Negative control 0.0 ± 0.0	N/A	N/A	N/A

empirically fit mortality and susceptibility functions (equations S2 and table S3) and compared this incomplete immunity model with one that still allows for higher host mortality and transmission for higher-virulence strains (*15*) but does not include immune protection from prior exposure. We applied an adaptive dynamics approach by simulating the invasion of a new mutant into a population with an endemic resident pathogen strain for a range of virulence levels of both the resident and invading strains (see supplementary materials). In the resulting pairwise invasibility plots (PIPs), the protective effects of host immunity shifted the pathogen's evolutionarily stable strategy to a virulence level almost twice as high as that in a model with no host protection (Fig. 3). This result, which is robust even when additional protective effects of strain homology are added to the model (fig. S3), is driven by two distinct mechanisms favoring greater virulence. First, because disease-induced mortality is a key cost of virulence, higher-virulence strains benefit the most from the reduction in host mortality (Fig. 2A) generated by incomplete immune protection. Second, the stronger protection provided by higher-virulence strains reduces the pool of previously infected hosts available for reinfection by lower-virulence strains (Fig. 2B). We evaluated the general applicability of our two-strain model to other study systems by conducting a numerical and analytical global sensitivity analysis. We found that a stronger relationship between virulence and immune protection enhances the competitive advantage of the more-virulent strain (tables S4 and S5). Thus, incomplete immunity should favor the evolution of greater virulence in any system in which higher-virulence strains generate a stronger protective effect against reinfection than lower-virulence strains (tables S4 and S5).

Our results, combined with high *M. gallisepticum* prevalence (*8*, *16*, *17*) and recovery rates (*6*) in the wild, indicate that host immunity plays a prominent role in the evolution of increased *M. gallisepticum* virulence in nature (*10*). Additional evolutionary processes, such as selection favoring higher transmission rates for more-virulent strains (*15*), may also contribute to the observed increases in virulence. Adaptation to a novel host is unlikely to explain the virulence increases on both coasts, as the more evolutionarily derived California strains of *M. gallisepticum* have lower virulence than eastern strains (Tables 1 and 2) (*10*, *18*). Although evolution of host resistance can lead to increased virulence (*19*) and may play a role in this system (*20*), there was no evidence of evolved host resistance in two studies conducted after the detected increases in virulence (*21*, *22*). Thus, our results suggest that *M. gallisepticum* virulence increased in both eastern and western populations as hosts with incomplete immunity became more common in each population.

The effects of incomplete immunity described here are arguably a specific case of a broader phenomenon whereby increased virulence is favored by quantitative host variation in susceptibility, whether due to host genetic variation (*23*), imperfect vaccines (*5*), or innate immune priming

(*24*). Disentangling differences in susceptibility versus infectiousness is a difficult problem in any system, and we do not attempt to do so here. Our model attributed differences in infection rates to what we term "host susceptibility" (η in equations S2), as infection occurred through inoculation with equal pathogen doses. It is likely that lower pathogen loads in hosts with prior exposure (Fig. 1B) will also lead to lower infectiousness and thus less transmission by those individuals. Although our models allowed for higher transmission rates for more-virulent strains in both scenarios (Fig. 3), we did not vary transmission rates with host prior exposure. Further, although our data support reduced infection length with prior exposure (fig. S2), we were unable to robustly quantify this effect and thus assumed equal infection lengths. Thus, our model is likely conservative, as selection for higher virulence should be even stronger if infectiousness and infection length also vary with host prior exposure.

Previous studies argue for great care in designing vaccines because incomplete protection can select for increased virulence in the targeted pathogens (*1*, *3*, *5*). Our study suggests that pathogens can readily evolve toward higher virulence due simply to the imperfect nature of host immune memory, whether via adaptive or innate responses (*14*). Despite historical focus on the small subset of pathogens that confer complete and lifelong immunity, incomplete immunity following infection is widespread in humans (*25*–*28*) and other animals (*29*–*32*). Because there are likely many systems where more-virulent pathogen strains stimulate stronger immune responses and thus provide stronger protection (*33*, *34*), the fitness advantages to higher virulence demonstrated here are relevant for a range of host and pathogen systems, including humans. Overall, our results show that the same immune systems that evolved to protect hosts from infection can drive the evolution of more-harmful pathogens in nature.

REFERENCES AND NOTES

1. S. Gandon, M. J. Mackinnon, S. Nee, A. F. Read, *Nature* **414**, 751–756 (2001).
2. M. J. Mackinnon, S. Gandon, A. F. Read, *Vaccine* **26** (suppl. 3), C42–C52 (2008).
3. P. D. Williams, T. Day, *Mol. Ecol.* **17**, 485–499 (2008).
4. V. C. Barclay *et al*., *PLOS Biol.* **10**, e1001368 (2012).
5. A. F. Read *et al*., *PLOS Biol.* **13**, e1002198 (2015).
6. C. R. Faustino *et al*., *J. Anim. Ecol.* **73**, 651–669 (2004).
7. J. S. Adelman, C. Mayer, D. M. Hawley, *J. Avian Biol.* **48**, 519–528 (2017).
8. A. A. Dhondt *et al*., *EcoHealth* **3**, 95 (2006).
9. A. A. Dhondt, K. V. Dhondt, D. M. Hawley, C. S. Jennelle, *Avian Pathol*. **36**, 205–208 (2007).
10. D. M. Hawley *et al*., *PLOS Biol.* **11**, e1001570 (2013).
11. A. A. Dhondt, K. V. Dhondt, W. M. Hochachka, D. H. Ley, D. M. Hawley, *Avian Dis*. **61**, 437–441 (2017).
12. A. E. Leon, D. M. Hawley, *EcoHealth* **14**, 793–804 (2017).
13. K. V. Sydenstricker, A. A. Dhondt, D. H. Ley, G. V. Kollias, *J. Wildl. Dis.* **41**, 326–333 (2005).
14. M. G. Netea, J. Quintin, J. W. M. van der Meer, *Cell Host Microbe* **9**, 355–361 (2011).
15. P. D. Williams, A. P. Dobson, K. V. Dhondt, D. M. Hawley, A. A. Dhondt, *J. Evol. Biol.* **27**, 1271–1278 (2014).
16. P. M. Nolan, G. E. Hill, A. M. Stoehr, *Proc. R. Soc. London Ser. B* **265**, 961–965 (1998).
17. C. S. Jennelle, E. G. Cooch, M. J. Conroy, J. C. Senar, *Ecol. Appl.* **17**, 154–167 (2007).
18. W. M. Hochachka *et al*., *Proc. R. Soc. London Ser. B* **280**, 20131068 (2013).
19. F. Fenner, F. N. Ratcliffe, *Myxomatosis* (Cambridge Univ. Press, Cambridge, 1965).
20. C. Bonneaud *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7866–7871 (2011).
21. D. M. Hawley *et al*., *J. Evol. Biol.* **23**, 1680–1688 (2010).
22. J. S. Adelman, L. Kirkpatrick, J. L. Grodio, D. M. Hawley, *Am. Nat.* **181**, 674–689 (2013).
23. C. E. L. Delmas *et al*., *Evol. Appl.* **9**, 709–725 (2016).
24. A. T. Tate, *Oikos* **126**, 350–360 (2017).
25. M. G. M. Gomes, A. Margheri, G. F. Medley, C. Rebelo, *J. Math. Biol.* **51**, 414–430 (2005).
26. W. M. Geisler, S. Y. Lensing, C. G. Press, E. W. Hook 3rd, *J. Infect. Dis.* **207**, 1850–1856 (2013).
27. S. Hall *et al*., *Pediatrics* **109**, 1068–1073 (2002).
28. X. Castellsagué *et al*., *J. Infect. Dis.* **210**, 517–534 (2014).
29. R. Casais *et al*., *Vet. Parasitol.* **203**, 173–183 (2014).
30. M. C. De Jong, W. H. van der Poel, J. A. Kramps, A. Brand, J. T. van Oirschot, *Am. J. Vet. Res.* **57**, 628–633 (1996).
31. A. Sabó, D. Blaškovic, *Acta Virol.* **14**, 17–24 (1970).
32. J. L. Schulman, E. D. Kilbourne, *J. Bacteriol.* **89**, 170–174 (1965).
33. M. D. Moody, C. M. Downs, *J. Bacteriol.* **70**, 297–304 (1955).
34. M. P. Davenport, G. T. Belz, R. M. Ribeiro, *Trends Immunol.* **30**, 61–66 (2009).
35. A. E. Fleming-Davies, V. Dukic, V. Andreasen, G. Dwyer, *Ecol. Lett.* **18**, 1252–1261 (2015).

ACKNOWLEDGMENTS

This work is dedicated to the memory of Paul Williams, who inspired us all with his insights and dedication. We thank L. Kirkpatrick, J. Adelman, S. Moyers, and numerous undergraduates for technical assistance. D. J. Páez and three anonymous reviewers provided helpful comments on the manuscript. This work was funded through NIH grant 5R01GM105245 to D.M.H. under the NIH-NSF-USDA Ecology and Evolution of Infectious Diseases program. Birds were captured under permits from the Virginia Department of Game and Inland Fisheries (044569) and the U.S. Fish and Wildlife Service (MB158404-1). All data and code to understand and assess the conclusions of this research are included in the main text or are available via Dryad Digital Repository (doi:10.5061/dryad.435h5).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1030/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S4
Tables S1 to S5
References (*36*–*49*)

30 June 2017; accepted 12 January 2018
10.1126/science.aao2140

ADAPTATION

Winter color polymorphisms identify global hot spots for evolutionary rescue from climate change

L. Scott Mills,[1,2]*† Eugenia V. Bragina,[2]† Alexander V. Kumar,[2,3] Marketa Zimova,[2,3] Diana J. R. Lafferty,[2,3] Jennifer Feltner,[2,3] Brandon M. Davis,[2,3] Klaus Hackländer,[2,4] Paulo C. Alves,[3,5,6] Jeffrey M. Good,[7] José Melo-Ferreira,[5,6] Andreas Dietz,[8] Alexei V. Abramov,[9] Natalia Lopatina,[10] Kairsten Fay[2]

Maintenance of biodiversity in a rapidly changing climate will depend on the efficacy of evolutionary rescue, whereby population declines due to abrupt environmental change are reversed by shifts in genetically driven adaptive traits. However, a lack of traits known to be under direct selection by anthropogenic climate change has limited the incorporation of evolutionary processes into global conservation efforts. In 21 vertebrate species, some individuals undergo a seasonal color molt from summer brown to winter white as camouflage against snow, whereas other individuals remain brown. Seasonal snow duration is decreasing globally, and fitness is lower for winter white animals on snowless backgrounds. Based on 2713 georeferenced samples of known winter coat color—from eight species across trophic levels—we identify environmentally driven clinal gradients in winter coat color, including polymorphic zones where winter brown and white morphs co-occur. These polymorphic zones, underrepresented by existing global protected area networks, indicate hot spots for evolutionary rescue in a changing climate.

The importance of evolution in fostering the persistence of species facing rapid environmental change is a fundamental tenet of biology that underlies the modern field of conservation biology (*1*–*3*). Despite the central role of evolution for maintaining biodiversity, criteria to facilitate adaptation by wild species remain largely absent from conservation planning (*4*, *5*). This is a particularly acute omission in a rapidly changing climate (*6*, *7*) where evolutionary rescue may reverse population declines via adaptive evolutionary change in phenotypes (*2*, *8*, *9*).

As a first step to demonstrate how evolutionary rescue might enter conservation planning for climate change, we describe a fitness-relevant trait that exhibits clines of locally adapted morphs shaped directly by climate. At least 21 bird and mammal species undergo photoperiod-induced seasonal coat color molts from brown to white in some portions of their range to maintain crypsis against seasonal snow presence or absence (Table 1). This seasonal phenological trait is confronting decreased seasonal snow cover duration, one of the most persistent and widespread signals of climate change (*10*, *11*). Field studies show that winter white animals mismatched against snowless ground suffer high fitness costs due to increased predator-caused mortality, which in the absence of evolutionary shifts would result in substantial population declines (*12*). In fact, coat color mismatch against decreased snow duration may have already contributed to range contractions for several species (*13*–*16*).

Although the seasonal brown-white-brown color trait is a classic polyphenism—whereby multiple morphs are produced by a single individual (*17*)—individuals in some populations molt to brown winter coats, thereby not undergoing the circannual color change. This intraspecific variation results in monomorphic winter white and brown populations but also in polymorphic populations that include sympatric winter white and brown color morphs. Importantly, this phenotypic variation is genetically determined: Latitudinal transplants, common garden, and breeding experiments with several seasonal color molting species have consistently showed minimal plasticity in the expression of winter phenotype and instead suggested a simple genetic basis involving one or a few major loci [e.g., (*18*–*22*)].

The enhanced standing phenotypic variation fostered by genetically based polymorphisms have long been linked to individual fitness and to potential for evolution to rescue populations from abrupt environmental change (*23*, *24*). Specifically, color polymorphisms have served as powerful models demonstrating evolution in nature, including iconic examples of evolutionary response to anthropogenic stressors (*25*–*28*). For the seasonal coat color trait, selection is expected to act on all winter color morphs based on local snow duration, but evolutionary rescue to changing climate should be enhanced by polymorphic regions where both brown and white winter morphs co-occur.

Here, we use a hierarchical approach across organismal scales (individual, population, and species) to spatially map geographic clines in winter coat color against local climate variables (*29*). We collated georeferenced descriptions of winter coat color from 2713 specimens spanning 60 countries across species ranges, with data sources including published accounts and specimens at 26 museums globally (table S1). From these georeferenced winter color morph samples, we built predictive models of winter color phenotypes across geographic ranges for eight mammal species that span trophic levels: four hare species and four carnivore species (three weasels and Arctic fox).

The response variable for our global generalized mixed model was the probability of an individual having a winter white coat, with species as a random effect and fixed effects including climate and landscape-level covariates (table S2). As expected for a trait under selection for crypsis against snow or bare ground, the most important covariates emerging from the global model were snow-cover duration and two climate variables affecting snow seasonality and transience. The probability of being white in winter (as opposed to brown) increased positively with snow duration and with seasonality (ranges of mean monthly temperatures), and negatively with isothermality (an index of snow transience).

Using the three environmental covariates identified in the best-fitting model, we created for each of the eight species a predictive range-wide map that assigned to each pixel a probability of an individual being white in winter (Fig. 1 and figs. S1 to S8). Based on fivefold cross validation, models fit georeferenced winter color morph data well (*29*). Across species, clinal gradients in winter color follow expected environmental gradients based on snow duration and ephemerality: Winter white morphs were more likely in regions with more persistent snowpack that tended to be more northern, higher elevation, and less maritime (Fig. 1). These results suggest that strong natural selection for camouflage against varying snow duration underlies phenotypic variation in winter color morphs across environmental gradients.

To identify hot spots that foster evolutionary rescue, we converted the continuous probabilities of individuals being winter white (versus brown) into polymorphic zones, using both a narrow (40% < Probability [winter white] < 60%) and broad (20% < Probability [winter white] < 80%) criteria. Depending on the species and criteria, polymorphic zones comprised 1 to 57% of a species' range (table S4). The species with the most widespread polymorphic zones (for narrow/broad criteria) are arctic fox (10%/57%), white-tailed jackrabbit

[1]Wildlife Biology Program and Office of the Vice President for Research and Creative Scholarship, University of Montana, Missoula, MT 59812, USA. [2]Fisheries, Wildlife, and Conservation Biology Program, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, USA. [3]Wildlife Biology Program, University of Montana, Missoula, MT 59812, USA. [4]Institute of Wildlife Biology and Game Management, BOKU, University of Natural Resources and Life Sciences, Vienna, Austria. [5]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal. [6]Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 4169-007 Porto, Portugal. [7]Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA. [8]German Aerospace Center, Earth Observation Center, German Remote Sensing Data Center, Oberpfaffenhofen, Wessling 82234, Germany. [9]Zoological Institute, Russian Academy of Science, Saint Petersburg 199034, Russia. [10]Institute of Systematics and Ecology of Animals SB RAS, Novosibirsk, 630091, Russia.
*Corresponding author. Email: scott.mills@umontana.edu
†These authors contributed equally to this work.

Table 1. The 21 vertebrate species known to exhibit seasonal coat color molt. The first eight species are those with sufficient sample sizes of georeferenced winter color phenotype to model range-wide distribution of color morphs. The other 13 species are those known to undergo seasonal coat color change in at least some populations. Species taxonomy follows the IUCN red list.

FAMILY/Species	Origin of sampled specimens		
	Museums	Literature, citizen science, trapping records, etc.	TOTAL
LEPORIDAE			
Snowshoe hare (*Lepus americanus*)	335	132	**467**
White-tailed jackrabbit (*Lepus townsendii*)	130	14	**144**
Mountain hare (*Lepus timidus*)	149	74	**223**
Japanese hare (*Lepus brachyurus*)	8	54	**62**
MUSTELIDAE			
Short-tailed weasel/stoat/ermine (*Mustela erminea*)	623	32	**655**
Long-tailed weasel (*Mustela frenata*)	444	36	**480**
Least weasel (*Mustela nivalis*)	606	30	**636**
CANIDAE			
Arctic fox (*Vulpes lagopus*)	26	20	**46**
OVERALL SAMPLE SIZE:	**2321**	**392**	**2713**

OTHER KNOWN COLOR CHANGING SPECIES

MURIDAE: Siberian (Djungarian) hamster (*Phodopus sungorus*); Collared lemming (*Dicrostonyx groenlandicus*); Wrangel Island collared lemming (*Dicrostonyx vinogradovi*); Palearctic collared lemming (*Dicrostonyx torquatus*); Ungava collared lemming (*Dicrostonyx hudsonius*); Richardson's collared lemming (*Dicrostonyx richardsoni*); Nelson's collared lemming (*Dicrostonyx nelsoni*); Ogilvie mountains collared lemming (*Dicrostonyx nunatakensis*)

LEPORIDAE: Arctic hare (*Lepus arcticus*), Alaskan hare (*Lepus othus*)

TETRAONIDAE: Rock ptarmigan (*Lagopus muta*); White-tailed ptarmigan (*Lagopus leucurus*); Willow ptarmigan (*Lagopus lagopus*)

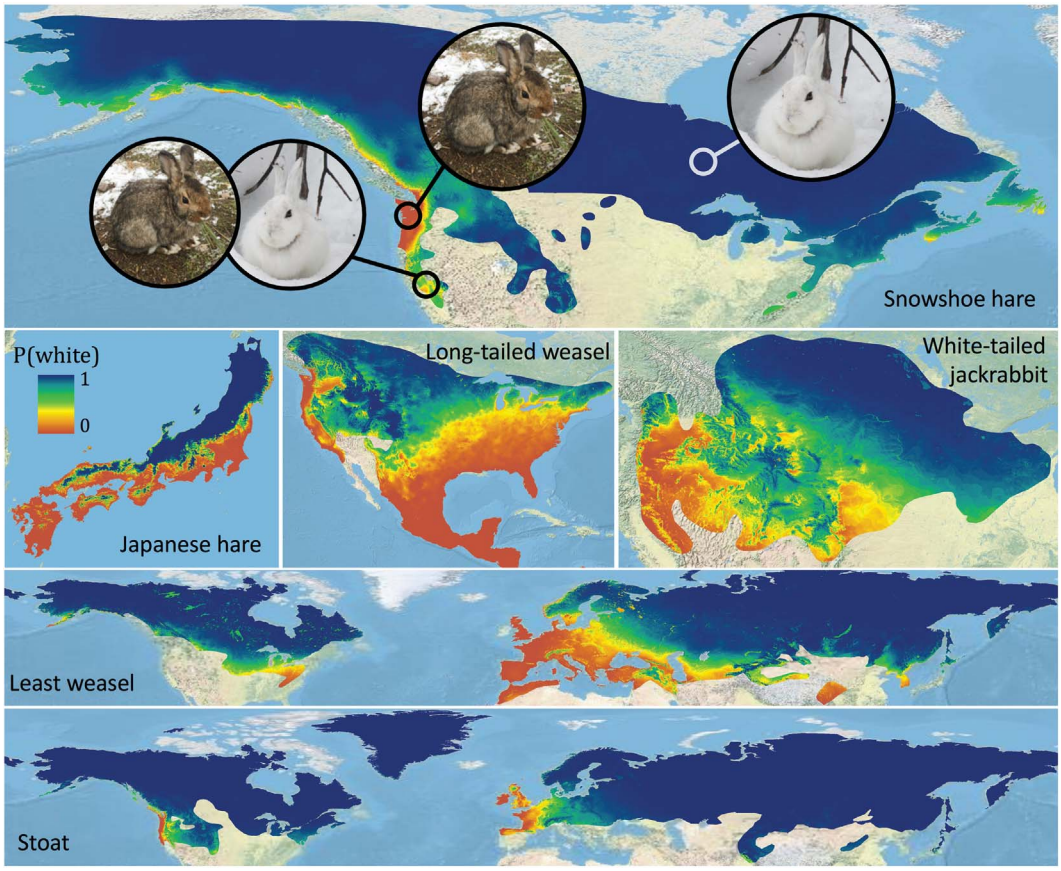


Fig. 1. Clinal variation in winter color phenotypes for six mammal species. Colder colors (e.g., blue) indicate higher probability of winter white morphs (denoted by photo of a winter white snowshoe hare), warmer colors (e.g., orange) indicate higher probability of winter brown morphs (denoted by brown snowshoe hare), and greenish/yellow colors indicate polymorphic populations (see figs. S1 to S8 for larger versions of these maps and for maps of Arctic fox and mountain hare). [Photo credits: L. S. Mills research archives]

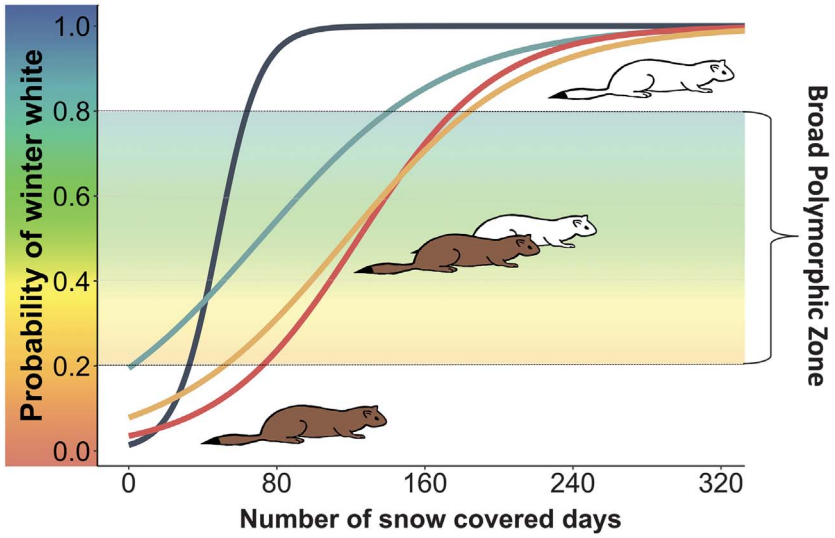


Fig. 2. Change in probability of being winter white as snow duration changes for four molting species. Species are Japanese hare, dark blue; white-tailed jackrabbit, light blue; least weasel, yellow; and long-tailed weasel, red. The central colored area with both winter white and brown animals represents our broadly defined polymorphic zone (i.e., 20% < P[winter white] < 80%).

Fig. 3. Regions with polymorphisms in winter coat color for multiple species. (**A** to **D**) Where polymorphic zones overlap for two (red) or three (brown) species, derived from predictive maps for eight species (see Fig. 1 and figs. S1 to S8). Polymorphic zones defined broadly as 20% < P[winter white] < 80% in (A) North America and (B) Eurasia. (C) Polymorphic zones defined more narrowly as 40% < P[winter white] < 60%; found only in (A) North America and (D) Great Britain. (**E**) Example of camouflage mismatch (least weasel). In polymorphic zones, as snow duration decreases, mismatched winter white morphs would be selected against in favor of the sympatric winter brown morphs. [Photo credit: Karol Zub]

(13%/43%), and long-tailed weasel (9%/33%). Mountain hares have the most restricted polymorphic zone (1%/2%).

Given that the clinal gradient of winter color represents fine-tuned adaptation to local snow conditions, how must winter phenotypes shift to adaptively track projected reductions in snow duration? Based on our model, we plotted the current probabilities of being white in winter against snow duration for four species (table S6) to characterize "optimal" winter coat color as shaped by past selection (Fig. 2). Depending on the species and snow duration, a plausible reduction of 30 to 50 days of seasonal snow cover during this century (*30*) would require many winter white populations to become polymorphic and polymorphic populations to become winter brown to maintain optimal winter coat colors.

Next, we combined the polymorphic zones of the eight species to identify regions with multispecies polymorphic zones (Fig. 3). Although under the broad criteria, two or more species shared putative polymorphic zones across much of the Northern Hemisphere (Fig. 3, A and B), narrow criteria multispecies polymorphic zones were limited to a few regions in North America (Fig. 3C) and Great Britain (Fig. 3D).

Polymorphic zones within and across these eight species ranges identify regions that currently hold disproportionately high potential to initiate evolutionary rescue from camouflage mismatch in this fitness-relevant trait affected by climate change. In addition to being hot spots for in situ evolutionary rescue, these areas may also facilitate gene flow of adaptive alleles to monomorphic populations (*31*, *32*).

Although protected areas cover 13% of the world's terrestrial area (*33*), multispecies polymorphic zones are poorly represented by existing protected areas (table S5). Even under our broad criteria, only 5% of multispecies polymorphic zones occur in the most strict protected areas described by the International Union for Conservation of Nature (IUCN) [categories I and II (*34*)]; all six IUCN categories of protected areas combined embrace only 10% of multispecies polymorphic zones (for the narrowly defined polymorphic zones; 4% fall in "strict" and 7% in "all") (table S5).

The broad geographic ranges of color molting species, and their roles as flagships and strongly interacting predators and prey, amplify the value of understanding how climate-mediated evolution may foster their persistence in the face of climate change. Failed adaptation by these species could have indirect impacts that reverberate through their ecosystems. Further, because the codistributed species that make up the multispecies polymorphic zones represent both predators (e.g., weasels and Arctic fox) and prey (e.g., hares), differential molt responses in different species could exacerbate fitness costs and create cascading coevolutionary outcomes.

Mismatch in seasonal coat color provides a visual metaphor for how climate change may affect biodiversity, and regions of sympatric winter color polymorphisms identify multispecies hot spots for evolutionary rescue in the face

of reduced snow duration. Our framework to identify zones of enhanced potential to initiate evolutionary rescue from climate change could be applied to polymorphisms in other morphological, physiological, or behavioral traits affected by climate change. Identification of hot spots for evolutionary rescue provides novel opportunities to integrate evolutionary processes to conservation planning in a changing climate.

REFERENCES AND NOTES

1. J. Huxley, *Heredity* **9**, 1–51 (1955).
2. E. Vander Wal, D. Garant, M. Festa-Bianchet, F. Pelletier, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120090 (2013).
3. M. E. Soulé, O. H. Frankel, *Conservation and Evolution* (Cambridge Univ. Press, Cambridge, 1981).
4. IUCN, Standard for the identification of key biodiversity areas, Version 1 (IUCN, Gland, Switzerland, 2015).
5. C. R. Margules, R. L. Pressey, *Nature* **405**, 243–253 (2000).
6. S. Lavergne, M. E. K. Evans, I. J. Burfield, F. Jiguet, W. Thuiller, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120091 (2013).
7. T. P. Dawson, S. T. Jackson, J. I. House, I. C. Prentice, G. M. Mace, *Science* **332**, 53–58 (2011).
8. A. Gonzalez, O. Ronce, R. Ferriere, M. E. Hochberg, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120404 (2013).
9. S. M. Carlson, C. J. Cunningham, P. A. H. Westley, *Trends Ecol. Evol.* **29**, 521–530 (2014).
10. N. S. Diffenbaugh, C. B. Field, *Science* **341**, 486–492 (2013).
11. K. E. Kunkel *et al.*, *Curr. Clim. Change Rep.* **2**, 65–73 (2016).
12. M. Zimova, L. S. Mills, J. J. Nowak, *Ecol. Lett.* **19**, 299–307 (2016).
13. S. M. Sultaire *et al.*, *Proc. R. Soc. London Ser. B* **283**, 20153104 (2016).
14. D. R. Diefenbach, S. L. Rathbun, J. K. Vreeland, D. Grove, W. J. Kanapaux, *Northeast. Nat. (Steuben)* **23**, 229–248 (2016).
15. S. Imperio, R. Bionda, R. Viterbi, A. Provenzale, *PLOS ONE* **8**, e81598 (2013).
16. S. Pedersen, M. Odden, H. C. Pedersen, *Ecosphere* **8**, 1234 (2017).
17. E. Mayr, *Animal Species and Evolution* (The Belknap Press of Harvard University Press, Cambridge, MA, 1963).
18. E. R. Hall, *Univ. Kans. Publ. Mus. Nat. Hist.* **4**, 1–466 (1951).
19. R. M. Hansen, G. D. Bear, *J. Mammal.* **44**, 420–422 (1963).
20. M. S. Ferreira *et al.*, *Mol. Ecol.* **26**, 4173–4185 (2017).
21. D. I. Våge *et al.*, *Peptides* **26**, 1814–1817 (2005).
22. A. Bergengren, *Viltrevy* **6**, 380–460 (1969).
23. A. Forsman, *Mol. Ecol.* **25**, 2693–2698 (2016).
24. L. Wennersten, A. Forsman, *Biol. Rev. Camb. Philos. Soc.* **87**, 756–767 (2012).
25. T. Caro, *Bioscience* **55**, 125–136 (2005).
26. M. Stevens, S. Merilaita, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 423–427 (2009).
27. E. B. Ford, *Biol. Rev. Camb. Philos. Soc.* **20**, 73–88 (1945).
28. I. J. Saccheri, F. Rousset, P. C. Watts, P. M. Brakefield, L. M. Cook, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 16212–16217 (2008).
29. Materials and methods are available as supplementary materials.).
30. L. S. Mills *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 7360–7365 (2013).
31. L. Hannah *et al.*, *Trends Ecol. Evol.* **29**, 390–397 (2014).
32. A. Hampe, A. S. Jump, *Annu. Rev. Ecol. Evol. Syst.* **42**, 313–333 (2011).
33. B. Bertzky *et al.*, Protected Plant Report 2012: Tracking progress towards global targets for protected areas (IUCN, Gland, Switzerland and UNEP-WCMC, Cambridge, UK, 2012).
34. A. Pfaff, F. Santiago-Ávila, L. Joppa, *Forests* **8**, 17 (2016).

ACKNOWLEDGMENTS

We thank S. T. Giery, D. Emlen, and four anonymous reviewers for insightful comments on the manuscript. We also thank museum curators (table S1) and substantial sample contributions from K. Zub, K. Garrison, T. Berry, J. Dines, E. M. Fouts, K. Nicholson, D. Oehlschlager, C. Opitz, A. M. Riedel, B. Wommack, S. Lado, M. Jones, J. Ivan, J. Whittington, and A. Angerbjörn. **Funding:** L.S.M.: National Science Foundation Division of Environmental Biology Grant 0841884; A.V.K.: National Science Foundation Graduate Research Fellowship under grant no. DGE-1252376; P.C.A.: FLAD, Luso-American Development Foundation; M.Z.: The Department of the Interior Southeast Climate Science Center Global Change Fellowship (Cooperative Agreement No. G10AC00624); J.M.-F.: Fundação para a Ciência e a Tecnologia Grants "CHANGE" PTDC/BIA-EVF/1624/2014 (Portuguese national funds) and IF/00033/2014/CP1256/CT0005 Investigador Fundação para a Ciência e a Tecnologia (Programa Operacional Potencial Humano-Quadro de Referência Estratégica Nacional funds from European Social Fund and Ministério da Ciência, Tecnologia e Ensino Superior), and SYNTHESYS grant SE-TAF-4695 (European Union Seventh Framework Programme; agreement 226506) to access the Swedish Museum of Natural History; J.M.G.: National Science Foundation EPSCoR grant 1736249; N.L.: Russian Foundation for Basic Research (17-04-00269 A). **Author contributions:** L.S.M. conceived the idea, helped with analysis, and led the writing. E.B., along with A.V.K., led the data analysis, contributed conceptually, and helped with writing. M.Z. helped with data analysis, figures, writing, and contributed conceptually. D.J.R.L., J.F., B.M.D., K.H., P.C.A., J.M.G., and J.M.-F. contributed conceptually and with writing. A.D., A.V.A., N.L., and K.F. helped provide key input data. All authors read and approved the submitted manuscript. **Competing interests:** None declared. **Data and materials availability:** The data reported in the paper are available in the supplementary materials and archived in the Dryad Digital Repository at https://doi.org/10.5061/dryad.8m0p1. Global SnowPack data are available as Web Map Service at DLR (German Aerospace Center) Geoservice, https://geoservice.dlr.de.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1033/suppl/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S6
References (*35*–*46*)

1 June 2017; accepted 2 January 2018
Published online 15 February 2018
10.1126/science.aan8097

IMMUNE ENGINEERING

Selective targeting of engineered T cells using orthogonal IL-2 cytokine-receptor complexes

Jonathan T. Sockolosky,[1,2] Eleonora Trotta,[3]* Giulia Parisi,[4]* Lora Picton,[1] Leon L. Su,[1] Alan C. Le,[5] Akanksha Chhabra,[5] Stephanie L. Silveria,[3] Benson M. George,[2,5,6] Indigo C. King,[7] Matthew R. Tiffany,[8] Kevin Jude,[1] Leah V. Sibener,[1,9] David Baker,[7] Judith A. Shizuru,[5] Antoni Ribas,[4,10] Jeffrey A. Bluestone,[3,10] K. Christopher Garcia[1,2,10,11]†

Interleukin-2 (IL-2) is a cytokine required for effector T cell expansion, survival, and function, especially for engineered T cells in adoptive cell immunotherapy, but its pleiotropy leads to simultaneous stimulation and suppression of immune responses as well as systemic toxicity, limiting its therapeutic use. We engineered IL-2 cytokine-receptor orthogonal (*ortho*) pairs that interact with one another, transmitting native IL-2 signals, but do not interact with their natural cytokine and receptor counterparts. Introduction of *ortho*IL-2Rβ into T cells enabled the selective cellular targeting of *ortho*IL-2 to engineered CD4$^+$ and CD8$^+$ T cells in vitro and in vivo, with limited off-target effects and negligible toxicity. *Ortho*IL-2 pairs were efficacious in a preclinical mouse cancer model of adoptive cell therapy and may therefore represent a synthetic approach to achieving selective potentiation of engineered cells.

Adoptive transfer of tumor-reactive T cells has evolved into a clinically useful therapy capable of inducing antitumor immunity in patients (*1*, *2*). However, the broad application of adoptive T cell transfer (ACT) therapies to treat cancer has several limitations, including the production of sufficient quantities of cells for infusion and the failure of transferred T cells to persist and remain functional in vivo. In the clinic, the concomitant administration of the T cell growth factor interleukin-2 (IL-2) improves the survival, function, and antitumor activity of transplanted T cells (*3*, *4*). However, the use of IL-2 to potentiate ACT is complicated by the pleiotropic nature of IL-2, which induces both immune stimulatory and suppressive T cell responses as well as potentially severe toxicities (*5*). This is governed by the interaction between IL-2 and the IL-2 receptor (IL-2R), which consists of α, β, and γ subunits (*6*). IL-2Rβ and the common γ-chain (IL-2Rγ) together form the signaling dimer and bind IL-2 with moderate affinity, whereas IL-2Rα (CD25) does not signal but increases the affinity of IL-2 for the binary (βγ) IL-2 receptor to sensitize T cells to low concentrations of IL-2.

The activity of IL-2 as an adjuvant to ACT is dependent on the balance between activation of transplanted and endogenous T cell subsets bearing natural IL-2 receptors, as well as host responses that cause dose-limiting toxicities. Strategies to overcome these limitations could improve T cell immunotherapy (*7*, *8*). Recognizing the need for new approaches that afford precise targeting of IL-2–dependent functions to a specific cell type of interest, we devised a strategy to redirect the specificity of IL-2 toward adoptively transferred T cells. This method, based on receptor-ligand orthogonalization, uses a mutant IL-2 cytokine and mutant IL-2 receptor that bind specifically to one another but not to their wild-type counterparts (Fig. 1A).

We focused on the murine IL-2/IL-2Rβ interaction to enable in vivo characterization in syngeneic mouse models. The IL-2Rβ chain was chosen as the mutant receptor because the β chain is required for signal transduction and can bind IL-2 independently. We devised a two-step approach to engineer orthogonal IL-2/IL-2Rβ pairs informed by the crystal structure of the IL-2 high-affinity receptor complex (*6*) (Fig. 1B). First, point mutations of the IL-2Rβ chain were identified from inspection of the interface between IL-2 and IL-2Rβ that abrogated binding to wild-type IL-2 (Fig. 1, C to E). The IL-2Rβ hotspot residues His^{134} and Tyr^{135} make numerous contacts with IL-2 that contribute a majority of the binding free energy between IL-2 and IL-2Rβ (*6*) (Fig. 1E). A double mutant IL-2Rβ [His^{134} → Asp (H134D) and Tyr^{135} → Phe (Y135F)], referred to herein as *ortho*IL-2Rβ, lacked detectable binding to IL-2 (Fig. 1D), even in the presence of CD25 (fig. S1) (*7*, *9*).

Next, we used yeast display-based evolution to mutate, and thus remodel, the wild-type IL-2 interface region that was opposing (or facing the site of) the IL-2Rβ mutations in the crystal structure, in order to create a molecule that bound to *ortho*IL-2Rβ but not to wild-type IL-2Rβ. IL-2 residues in proximity to the *ortho*IL-2Rβ binding interface were randomly mutated and were chosen on the basis of a homology model of the mouse IL-2/IL-2Rβ complex (Fig. 1E) derived from the crystal structure of the human IL-2 receptor complex (*6*). A library of ~10^8 unique IL-2 mutants was displayed on the surface of yeast (fig. S2) and subjected to multiple rounds of both positive (against *ortho*IL-2Rβ) and negative (against IL-2Rβ) selection (figs. S2 and S3). This collection of yeast-displayed IL-2 mutants bound the *ortho*IL-2Rβ, but not wild-type IL-2Rβ, and retained CD25 binding (Fig. 1D). Sequencing of yeast clones from the evolved IL-2 libraries revealed a consensus set of mutations at IL-2 positions in close structural proximity to the *ortho*IL-2Rβ mutations (fig. S4). Interestingly, a Gln^{30} → Asn (Q30N) mutation was highly conserved across three independent mutant IL-2 yeast libraries, whereas all other IL-2 positions used a restricted but not specific mutational signature. We found that IL-2 mutations Q30N, Met^{33} → Val (M33V), and Asp^{34} → Leu or Met (D34L/M) appear to form a small nonpolar pocket to compensate for the IL-2Rβ Y135F mutation, whereas Gln^{36} → Thr, Ser, Lys, or Glu (Q36T/S/K/E) and Glu^{37} → Tyr or His (E37Y/H) mutations present a polar or charged surface to compensate for the IL-2Rβ H134D mutation (Fig. 1F).

Because of the affinity-enhancing effects of CD25 expression on the interaction of IL-2 with the binary (βγ) IL-2 receptor (*10*), IL-2 mutants with negligible binding to IL-2Rβ alone may still form a functional signaling complex on cells that also express CD25 (*8*). Therefore, we used a yeast-based functional screen to further triage IL-2 mutants that bound specifically to the *ortho*IL-2Rβ and signaled selectively on T cells that express the *ortho*IL-2Rβ (Fig. 1G and fig. S5), and produced recombinant forms of select IL-2 mutants (*ortho*IL-2) for characterization (figs. S6 to S8).

We focused our efforts on two *ortho*IL-2 mutants, 1G12 and 3A10. *Ortho*IL-2 1G12 and 3A10 share the consensus Q30N, M33V, and D34L mutations but differ at positions Glu^{29}, Gln^{36}, Glu^{37}, and Arg^{41} (Fig. 1I). *Ortho*IL-2 1G12 and 3A10 bound the *ortho*IL-2Rβ with an affinity comparable to that of the wild-type IL-2/IL-2Rβ interaction and displayed little to no detectable binding to wild-type IL-2Rβ (Fig. 1H and figs. S7 and S8) but differed in their ability to activate IL-2Rβ signaling in CD25-positive wild-type

[1]Departments of Molecular and Cellular Physiology and Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA. [2]Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA 94305, USA. [3]Diabetes Center and Department of Medicine, University of California, San Francisco, CA 94143, USA. [4]Division of Hematology-Oncology, Department of Medicine, David Geffen School of Medicine, and Jonsson Comprehensive Cancer Center, University of California, Los Angeles, CA 90095, USA. [5]Department of Blood and Marrow Transplantation, Institute for Stem Cell Biology and Regenerative Medicine, and Ludwig Center for Cancer Stem Cell Research and Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA. [6]Stanford Medical Scientist Training Program, Stanford University, Stanford, CA 94305, USA. [7]Department of Biochemistry, Howard Hughes Medical Institute, and Institute for Protein Design, University of Washington, Seattle, WA 98195, USA. [8]Department of Pediatrics and Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA. [9]Immunology Graduate Program, Stanford University School of Medicine, Stanford, CA 94305, USA. [10]Parker Institute for Cancer Immunotherapy, 1 Letterman Drive, Suite D3500, San Francisco, CA 94129, USA. [11]Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: kcgarcia@stanford.edu

and *ortho*IL-2Rβ T cells. Stimulation of *ortho*IL-2Rβ T cells (fig. S5B) with *ortho*IL-2 1G12 resulted in dose-dependent phosphorylation of STAT5 (pSTAT5), a hallmark of IL-2R signaling, with potency similar to that of wild-type IL-2, but also induced pSTAT5 on wild-type T cells, albeit with significantly reduced potency relative to IL-2 (Fig. 1, G and I, and fig. S6). By comparison, *ortho*IL-2 3A10 was specific for *ortho*IL-2Rβ T cells, but with a weaker potency relative to IL-2 (Fig. 1, G and I, and fig. S6). We speculated that *ortho*IL-2 1G12 activity on wild-type T cells is a consequence of weak residual binding to wild-type IL-2Rβ (fig. S7). Low-affinity interactions with IL-2Rβ alone are enhanced in the presence of CD25 (*8*). Indeed, *ortho*IL-2 1G12 exhibited binding to wild-type IL-2Rβ when first captured by CD25, with limited binding in the absence of CD25 (figs. S1 and S8). *Ortho*IL-2 3A10 did not bind appreciably to IL-2Rβ even in the presence of CD25, in agreement with its negligible biological activity on CD25-positive T cells. Interaction of *ortho*IL-2 1G12 and 3A10 with *ortho*IL-2Rβ was significantly enhanced in the presence of CD25, with apparent binding affinities of the ternary CD25/*ortho*IL-2Rβ/*ortho*IL-2 complex that correlate with their respective potency on *ortho*IL-2Rβ T cells (fig. S1).

In clinical ACT regimens, patient-derived T cells for ACT are expanded in IL-2 before re-infusion in order to obtain sufficient numbers of therapeutic cells with the desired genotype/phenotype (*2*). We explored the in vitro activity of *ortho*IL-2 on activated primary mouse $CD8^+$ T cells engineered to express the *ortho*IL-2Rβ and a yellow fluorescent protein (YFP) to distinguish modified (YFP^+) and unmodified (YFP^-) cells (Fig. 2A). The transcription factor STAT5 is phosphorylated upon IL-2 engagement with the IL-2R and translocates to the nucleus, where it promotes the proliferation and cell cycle progression of T cells (*11*). Wild-type IL-2 induced the phosphorylation of STAT5 (pSTAT5) in both wild-type and *ortho*IL-2Rβ $CD8^+$ T cells with similar potency and signaling amplitude, indicating functional signal transduction through the wild-type receptor but not *ortho*IL-2Rβ (Fig. 2B). By comparison, *ortho*IL-2 1G12 potently activated STAT5 on *ortho*IL-2Rβ–transduced T cells, with a potency increase by a factor of ~5 relative to wild-type T cells. *Ortho*IL-2 3A10 induced somewhat weaker, albeit selective pSTAT5 on *ortho*IL-2Rβ–expressing but not wild-type T cells (Fig. 2, B, D, and E). These results were consistent with the biased binding of the *ortho*IL-2s to the *ortho*IL-2Rβ, which translated into the selective or specific expansion of *ortho*IL-2Rβ T cells cultured ex vivo in *ortho*IL-2 1G12 or 3A10, respectively (Fig. 2, C and D). The *ortho*IL-2Rβ–transduced T cells cultured in saturating concentrations of *ortho*IL-2 3A10 became enriched to near homogeneity after 3 to 5 days (Fig. 2F).

IL-2 is indispensable for the development and function of regulatory T cells (T_{regs}) (*12*), which are sensitive to IL-2 as a result of constitutive expression of CD25 and require IL-2Rβ–dependent activation of STAT5 signaling for survival and function (*13*). Both *ortho*IL-2 1G12 and 3A10 retained specificity for T_{regs} modified to express the *ortho*IL-2Rβ, with potency similar to that on $CD8^+$ T cells (Fig. 2G and fig. S9, A and B). In addition to cells that naturally respond to IL-2, activation of *ortho*IL-2Rβ signaling pathways with *ortho*IL-2 could, in principle, be achieved in any cell type that also expresses the IL-2Rγ. Activated mouse B cells expressed the IL-2Rγ but lacked appreciable levels of IL-2Rβ (*14*, *15*) and were relatively insensitive to IL-2–dependent STAT5 activation

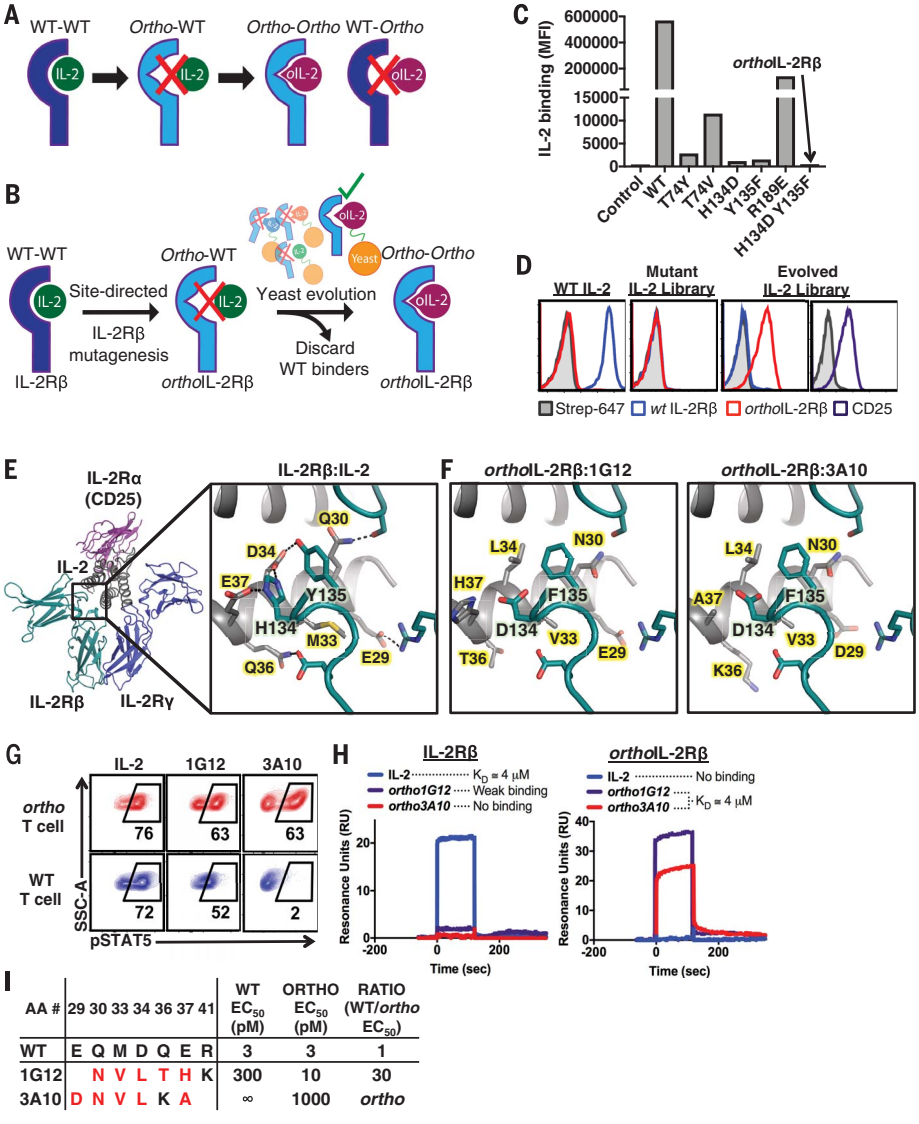


Fig. 1. Engineering and characterization of orthogonal IL-2 and IL-2R pairs. (**A**) Schematic overview of orthogonal IL-2/IL-2R pairs, consisting of a mutant IL-2 cytokine and mutant IL-2R that interact specifically with each other but do not cross-react with their wild-type counterparts. (**B**) Strategy used to engineer orthogonal IL-2/IL-2Rβ pairs. (**C**) Wild-type and mutant IL-2Rβ tetramer binding to wild-type IL-2 displayed on yeast by fluorescence-activated cell sorting. MFI, mean fluorescence intensity. Data are representative of two independent experiments. (**D**) Histograms of wild-type IL-2Rβ (blue), *ortho*IL-2Rβ (red), or CD25 (purple) binding to yeast-displayed wild-type IL-2, the naïve mutant IL-2 yeast library, or mutant IL-2 yeast clones after in vitro evolution. In vitro evolution of three independent mutant IL-2 yeast libraries (fig. S4) yielded similar results. (**E**) Homology model of the mouse IL-2/IL-2Rβ structure and the site I interface of IL-2 (gray) and contacts with IL-2Rβ His^{134} and Tyr^{135} (teal). Dashed lines indicate potential polar contacts. (**F**) Model of the *ortho*IL-2/*ortho*IL-2Rβ interactions. (**G**) Off-yeast pSTAT5 functional screen of IL-2 mutant activity on wild-type and *ortho*IL-2Rβ CTLL-2 T cells. (**H**) Representative surface plasmon resonance (SPR) sensograms of wild-type and *ortho*IL-2 binding to wild-type IL-2Rβ or *ortho*IL-2Rβ. Data are representative of two independent experiments. K_D, dissociation constant. (**I**) Sequences of wild-type (WT) IL-2, *ortho*IL-2 1G12, and *ortho*IL-2 3A10 and corresponding in vitro bioactivity (pSTAT5 EC_{50}) on wild-type and *ortho*IL-2Rβ CTLL-2 T cells. Amino acid codes: A, Ala; D, Asp; E, Glu; F, Phe; H, His; K, Lys; L, Leu; M, Met; N, Asn; Q, Gln; T, Thr; V, Val; Y, Tyr.

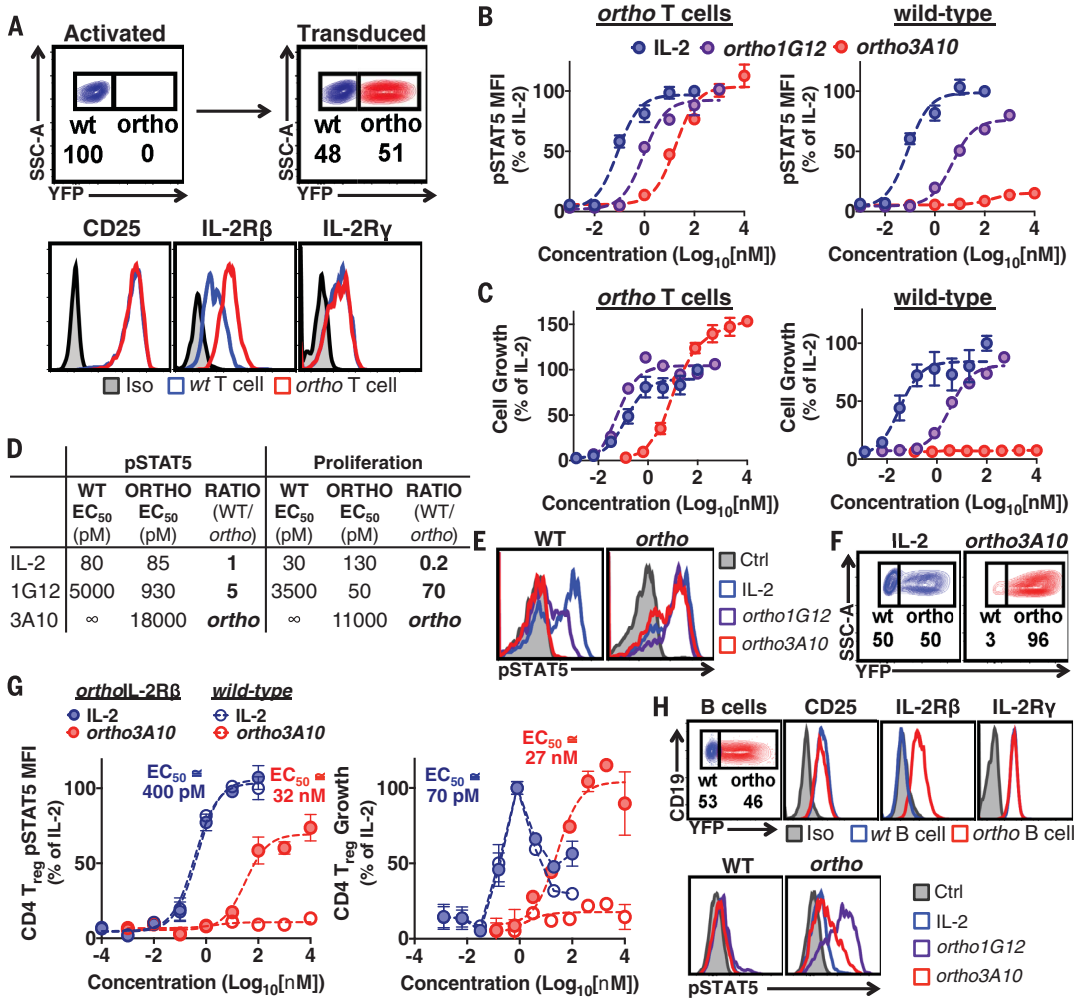


Fig. 2. *ortho*IL-2 signals through the *ortho*IL-2R expressed in primary mouse lymphocyte subsets, resulting in specific expansion of CD4 and CD8 T cells in vitro. (**A**) Flow cytometry data of mouse T cells transduced with the *ortho*IL-2Rβ and a YFP reporter (top panels) and associated cell surface levels of CD25, IL-2Rβ, and IL-2Rγ. (**B** to **F**) Dose-response curves of (B) STAT5 phosphorylation after 20 min of stimulation and (C) proliferation of wild-type (open circles) and *ortho*IL-2Rβ (solid circles) $CD8^+$ T cells cultured for 4 days in IL-2 or *ortho*IL-2; (D) table of respective pSTAT5 and proliferation EC_{50} from data in (B) and (C). Representative histograms of (E) STAT5 phosphorylation and (F) scatterplots of $CD8^+$ wild-type (YFP^-) and *ortho*IL-2Rβ (YFP^+) T cells expanded in IL-2. Data are means ± SD (n = 3 biological replicates). Dashed lines represent curves fit to a log (agonist) versus response (three parameters) model in Prism. (**G**) Dose-response curves of STAT5 phosphorylation (left) and proliferation (right) of wild-type and *ortho*IL-2Rβ $CD4^+$ T_{regs} cultured in IL-2 or *ortho*IL-2. Data are means ± SD (n = 3 biological replicates). (**H**) Representative histograms of primary mouse B cells transduced with the *ortho*IL-2Rβ and stimulated with the indicated cytokines for quantification of intracellular pSTAT5 as in fig. S9.

(Fig. 2H and fig. S9, E and F). Transduction of the *ortho*IL-Rβ into activated B cells rendered them responsive to *ortho*IL-2 (Fig. 2H and fig. S9, E and F), but with reduced potency and increased specificity relative to T cells. Specificity was due to the lack of appreciable wild-type IL-2Rβ on B cells (fig. S9E).

In a host with an intact immune system, adoptively transferred T cells must compete with host cells for survival signals such as IL-2 (*16*). However, unlike wild-type IL-2, there should be minimal competition from endogenous cells for *ortho*IL-2 consumption. Thus, we determined the in vivo activity of *ortho*IL-2 and *ortho*IL-2Rβ T cells in mice with intact immune systems. A mixture of wild-type and *ortho*IL-2Rβ $CD8^+$ T cells was adoptively transferred into wild-type mice, and the impact of IL-2 and *ortho*IL-2 administration on transplanted T cells and the host immune system was quantified (Fig. 3A). *Ortho*IL-2 1G12 significantly expanded $CD8^+$ T cells transduced with the *ortho*IL-2Rβ at doses equivalent to or lower than wild-type IL-2, which acted through the endogenous IL-2Rβ expressed in both wild-type and *ortho*IL-2Rβ T cells (Fig. 3B and fig. S10). The selectivity of *ortho*IL-2 1G12 for *ortho*IL-2Rβ T cells was dose-dependent, with increased activity on wild-type cells at increased dose amounts and/or frequency of treatment (Fig. 3, B and C, and figs. S10 to S12). These results were consistent with the in vitro selectivity of *ortho*IL-2 1G12, although it remained possible that *ortho*IL-2 1G12 signaling through the *ortho*IL-2Rβ could trigger endogenous IL-2 production by the *ortho*IL-2Rβ T cells, leading to indirect signaling through the wild-type IL-2R in cis or trans.

At high doses and twice-daily administration, *ortho*IL-2 3A10 resulted in the substantial expansion of *ortho*IL-2Rβ T cells with high specificity and no wild-type T cell expansion (Fig. 3, B and C, and figs. S11 and S12). This finding suggests that the effects of high-dose *ortho*IL-2 1G12 treatment were due not to induction of endogenous IL-2 by *ortho*IL-2Rβ $CD8^+$ T cells, but rather to low-level cross-reactivity with the wild-type IL-2Rβ by this molecule. The *ortho*IL-2 variants also promoted the in vivo expansion of *ortho*IL-2Rβ $CD4^+$ effector T cell (T_{eff}) (Fig. 3I and fig. S12) and *ortho*IL-2Rβ $CD4^+$ T_{reg} (fig. S9, C and D) cell subsets with specificity similar to that in $CD8^+$ T cells.

The two different *ortho*IL-2 variants exhibited specificities in vivo that mirrored their relative specificities in vitro. Despite its ability to activate wild-type IL-2Rβ signaling, albeit with about one order of magnitude less potency than *ortho*IL-2Rβ signaling, *ortho*IL-2 1G12 administration was relatively specific for *ortho*IL-2Rβ T cells in vivo (Fig. 3, B to H, and figs. S10 to S12). In mice treated twice daily with *ortho*IL-2 1G12 only, $CD4^+$ T_{regs} were elevated to a substantially lower degree than observed in IL-2–treated mice (Fig. 3F). However, the *ortho*IL-2 3A10 variant, consistent with the lack of wild-type IL-2Rβ signaling, had no detectable activity on host cell subset numbers (fig. S11) or expression of CD25, PD-1, and TIM-3, which are up-regulated by early or late IL-2R signaling (fig. S13).

To improve in vivo half-life and enable more convenient dosing, we fused IL-2 and *ortho*IL-2 to mouse serum albumin (*17*) (MSA), which has been shown to extend the half-life of mouse IL-2 from 5 hours to 50 hours (*18*). Fusion to MSA had little to no impact on IL-2– or *ortho*IL-2–dependent T cell proliferation in vitro (fig. S14); however, the in vivo activity was greatly enhanced. Fusion of

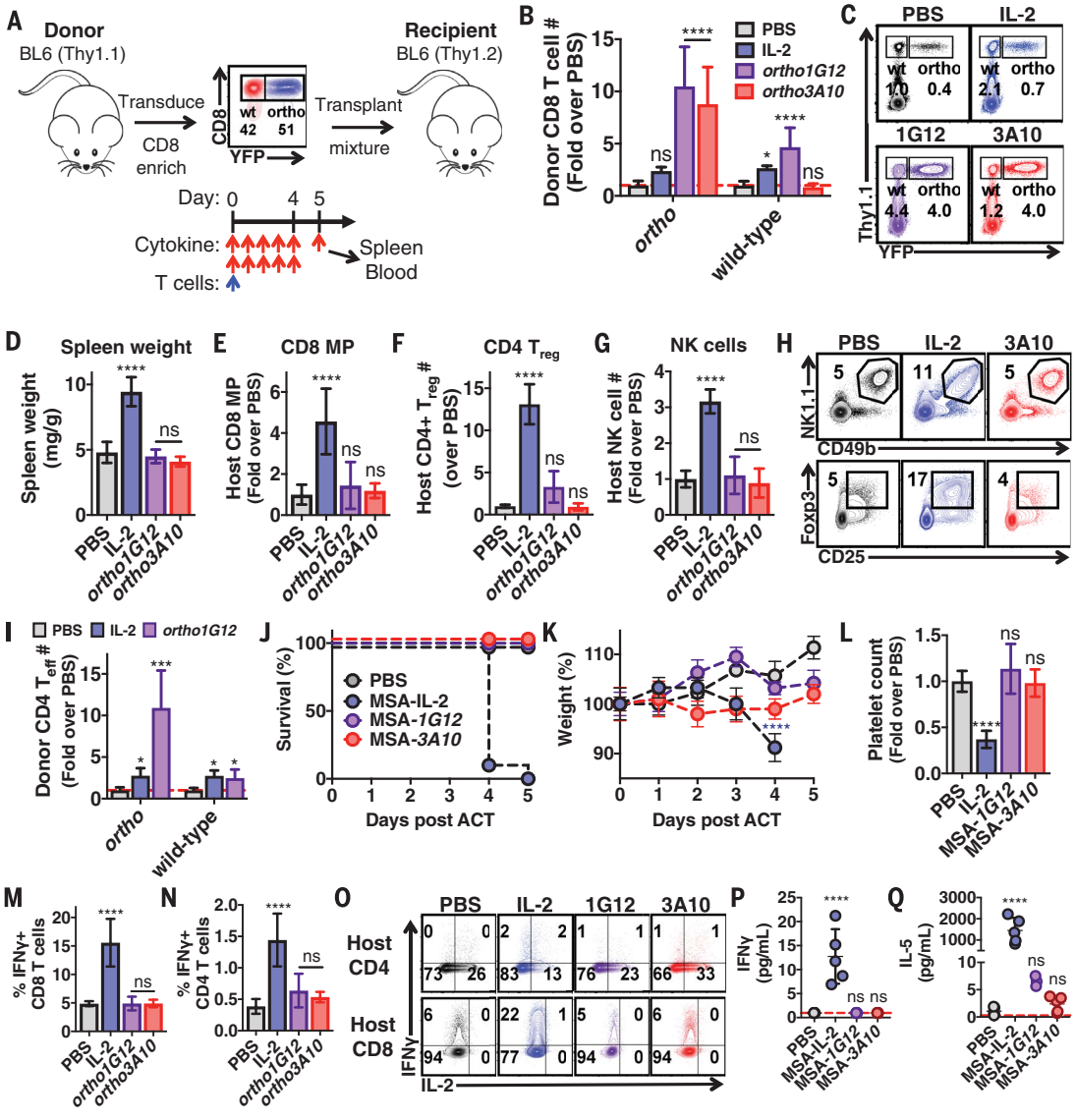


Fig. 3. *Ortho*IL-2 promotes the specific expansion of *ortho*IL-2Rβ–modified T cells in mice with negligible toxicity. (**A**) Schematic of the adoptive $CD8^+$ T cell transplant mouse model. (**B**) Quantification of donor wild-type and *ortho* $CD8^+$ T cells in the spleen of recipient mice treated twice daily with phosphate-buffered saline (PBS), IL-2 (250,000 IU/dose), *ortho*IL-2 1G12 (250,000 IU/dose), or *ortho*IL-2 3A10 (2,500,000 IU/dose). (**C**) Representative flow cytometry data quantified in (B) depicting donor ($Thy1.1^+$) wild-type (YFP^-) and *ortho*IL-2Rβ (YFP^+) $CD8^+$ T cells in the spleen of recipient mice. (**D**) Spleen weight of mice treated in (B) normalized to total body weight on day of killing. (**E** to **G**) Quantification of exogenous cytokine administration on host (E) $CD8^+$ memory phenotype T cell (MP, $CD44^+CD62L^+$), (F) $CD4^+$ T_{reg} ($CD25^+Foxp3^+$), and (G) natural killer (NK) cell ($CD3\text{-}NK1.1^+CD49b^+$) numbers in the spleen of mice treated in (A). (**H**) Representative flow cytometry data as quantified in (F) and (G). Data in (B) to (H) are means ± SD (n = 5 mice per group). *P < 0.05, ****P < 0.0001 [analysis of variance (ANOVA)]; ns, not significant. (**I**) Quantification of donor wild-type and *ortho*IL-2Rβ $CD4^+$ T_{eff} in the spleen of recipient mice treated once daily with PBS, IL-2 (250,000 IU/dose), or *ortho*IL-2 1G12 (1,000,000 IU/dose). Data are means ± SD and are representative of two independent experiments (n = 4 mice per group). *P < 0.05, ***P < 0.001 (ANOVA). (**J**) Survival of mice that received a mixture of wild-type and *ortho*IL-2Rβ $CD8^+$ T cells followed by daily administration of IL-2 or *ortho*IL-2 fused to MSA. All mice received a total of 250,000 IU/day of the respective MSA fusion protein on an IL-2 basis for 5 days. (**K**) Mouse body weight over time normalized to the group average on day 0 as treated in (J). (**L**) Platelet counts in peripheral blood on day 4 as treated in (J). Data in (J) to (L) are means ± SD (n = 5 mice per group). ****P < 0.0001 (ANOVA). (**M** to **O**) Quantification of cytokine administration on host (M) $CD8^+$ and (N) $CD4^+$ T cell production of IFN-γ upon ex vivo restimulation with phorbol 12-myristate 13-acetate (PMA) and ionomycin. (O) Representative flow cytometry data as quantified in (M) and (N). (**P** and **Q**) Serum (P) IFN-γ and (Q) IL-5 concentrations on day 7 in mice treated daily with PBS or with MSA–IL-2, MSA-1G12, or MSA-3A10 (each 25,000 IU/dose) for 7 days. Data are means ± SD (n = 5 mice per group). ****P < 0.0001 (ANOVA).

*ortho*IL-2 1G12 to MSA substantially increased its activity on cells that express the wild-type IL-2R relative to native *ortho*IL-2 1G12, leading to increased off-target effects and toxicity (fig. S15). However, the MSA–*ortho*IL-2 3A10 fusion protein retained exquisite specificity for *ortho*IL-2Rβ T cells (fig. S16).

One of the major limitations of IL-2 in the clinic is that IL-2 toxicity limits the use of high-dose IL-2 therapy for metastatic cancer and as an adjuvant to adoptive T cell therapy (*12*). IL-2 administered as a MSA fusion resulted in a number of dose-dependent and dose-accumulating toxicities that led to weight loss, restricted mobility, hypothermia, ruffled fur, hunched posture, splenomegaly, lymphomegaly, and death (Fig. 3, J to L, and figs. S15 to S18). In contrast, MSA–*ortho*IL-2 3A10 was nontoxic at all doses evaluated. MSA–*ortho*IL-2 3A10 activity was negligible on all IL-2–responsive host cell subsets evaluated.

In addition to its role as a proliferative cytokine, IL-2 is a potent effector cytokine capable of activating cytotoxic T cell functions and T cell inflammatory pathways (*19*). We determined the capacity of adoptively transferred *ortho*IL-2Rβ $CD8^+$ T cells to produce interferon-γ (IFN-γ) and cell surface levels of the immune inhibitory receptors PD-1 and TIM-3 after expansion in vivo with *ortho*IL-2. TIM-3 expression correlates with a highly dysfunctional $CD8^+$ T cell state, whereas PD-1 expression is associated with both T cell activation and exhaustion (*20*). *Ortho*IL-2Rβ T cells expanded in *ortho*IL-2 produced significantly more

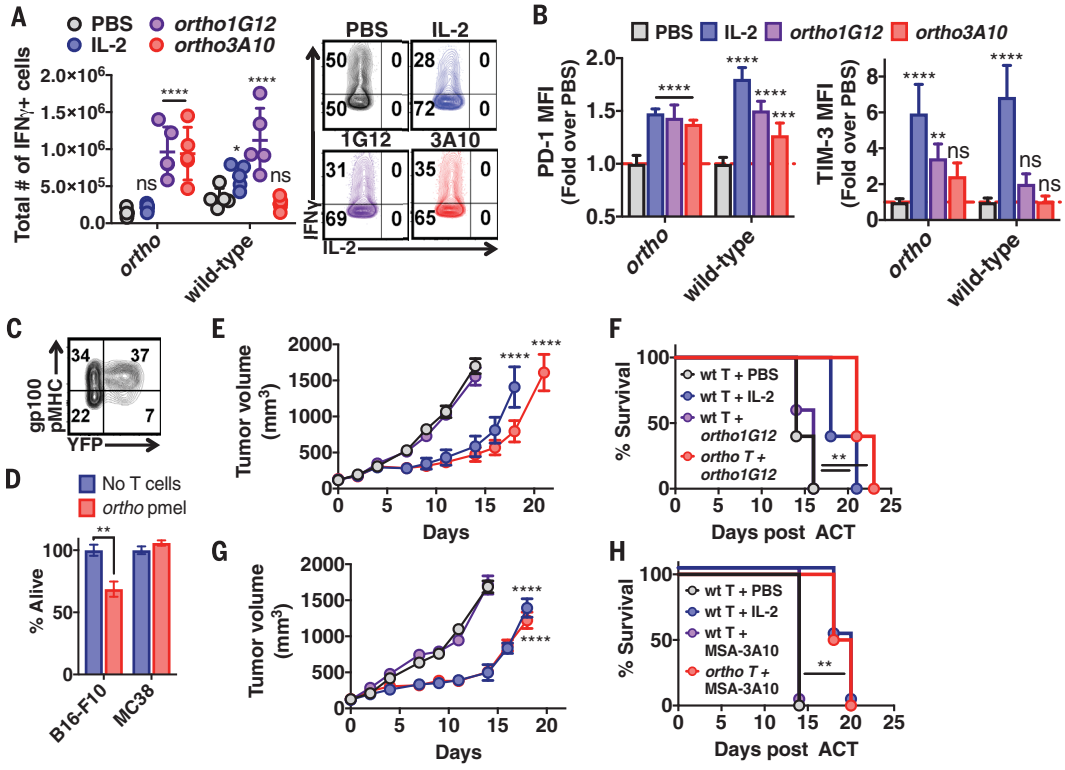


Fig. 4. *Ortho*IL-2–expanded T cells retain effector function and promote an antitumor response against syngeneic B16-F10 tumors in mice. (**A**) Quantification of total number of IFN-γ–positive wild-type or *ortho*IL-2Rβ $CD8^+$ T cells recovered from the spleen as treated in Fig. 3 (left) and representative flow cytometry data (right). (**B**) Cell surface expression levels of PD-1 (left) and TIM-3 (right) on wild-type and *ortho*IL-2Rβ $CD8^+$ T cells in the spleen after administration of the indicated cytokines. Data are means ± SD (n = 5 mice per group). *P < 0.05, ****P < 0.0001 (ANOVA). (**C**) gp100 pMHC tetramer staining of *ortho*IL-2Rβ–transduced pmel-1 transgenic $CD8^+$ T cells. (**D**) In vitro cytotoxicity of *ortho*IL-2Rβ pmel-1 transgenic T cells against antigen-positive (B16-F10) but not antigen-negative (MC38) tumor cells at a 20:1 (E:T) ratio. Data are means ± SD (n = 3 biological replicates). **P < 0.01 (Student t test). (**E** and **F**) Tumor growth (E) and survival (F) of C57BL/6J mice bearing subcutaneous B16-F10 tumors treated with wild-type (wt T) or *ortho*IL-2Rβ pmel-1 transgenic $CD8^+$ T cells (*ortho* T) and IL-2 or *ortho*IL-2 1G12. Data are means ± SEM (n = 5 mice per group). ****P < 0.0001 (two-way ANOVA) (E); **P < 0.01 (log-rank test) (F). (**G** and **H**) Tumor growth (G) and survival (H) of C57BL/6J mice bearing subcutaneous B16-F10 tumors treated with wild-type (wt T) or *ortho*IL-2Rβ pmel-1 transgenic $CD8^+$ T cells (*ortho* T) and IL-2 or *ortho*IL-2 3A10 fused to MSA. Data are means ± SEM (n = 4 mice per group). ****P < 0.0001 (two-way ANOVA) (G); **P < 0.01 (log-rank test) (H).

IFN-γ than IL-2–expanded cells (Fig. 4A). PD-1 levels were similar on *ortho*IL-2Rβ T cells from both IL-2– and *ortho*IL-2–treated mice (Fig. 4B). Interestingly, TIM-3 levels were significantly lower on *ortho*IL-2Rβ T cells from mice treated with *ortho*IL-2 relative to those treated with IL-2 (Fig. 4B).

The differential activity of *ortho*IL-2 on both T cell expansion and function may be due to increased bioavailability of *ortho*IL-2 for *ortho*IL-2Rβ T cells as the result of a reduced antigen sink or alternative host factors influenced by IL-2 but not *ortho*IL-2, which in turn may influence the function of transplanted T cells. For instance, IL-2 but not *ortho*IL-2 treatment increased host $CD4^+$ and $CD8^+$ T cell IFN-γ production upon ex vivo restimulation (Fig. 3, M to O) and increased the serum concentration of numerous inflammatory cytokines, including IFN-γ, IL-4, IL-5, IL-6, and IL-13 (Fig. 3, P and Q, and fig. S17). The ability to decouple direct IL-2 activity on transplanted T cells from indirect host bystander effects using *ortho*IL-2/IL-2R pairs may have important therapeutic implications.

To investigate prospective clinical applications of orthogonal IL-2/IL-2R pairs, we determined the efficacy of tumor-specific *ortho*IL-2Rβ T cells in the B16-F10 mouse model of melanoma. Transgenic pmel-1 T cell receptor (TCR) cells (pmel-1 T cells) express a high-affinity TCR that recognizes the B16-F10 specific ortholog of human gp100 (*19*), a self antigen overexpressed in human melanoma (Fig. 4, C and D). Adoptive transfer of pmel-1 T cells in combination with lymphocyte depletion and IL-2 administration can model ACT approaches to treat human cancer. Adoptive transfer of pmel-1 T cells accompanied by five daily injections of IL-2 significantly delayed tumor growth in mice and increased survival relative to mice treated only with T cells and saline (Fig. 4, E to G). Transfer of *ortho*IL-2Rβ pmel-1 T cells followed by treatment with native *ortho*IL-2 1G12 at a dose that had minimal activity on wild-type IL-2R cells (fig. S10) produced a significant tumor growth delay and survival advantage that mirrored the IL-2 treatment group (Fig. 4, E and F). Similar antitumor responses were observed in mice treated with *ortho*IL-2Rβ pmel-1 T cells and MSA–*ortho*IL-2 3A10 (Fig. 4, G and H). There was no therapeutic benefit of *ortho*IL-2 in mice that received wild-type pmel-1 T cells, indicating that *ortho*IL-2 activity is dependent on expression of the *ortho*IL-2Rβ in pmel-1 T cells.

Our results constitute an approach to redirect the specificity of IL-2 toward engineered T cells using orthogonal IL-2 cytokine-receptor pairs, which enables the selective expansion of desired T cell subsets in settings of adoptive cell therapy, but with limited off-target activity and negligible toxicity. Engineering orthogonal molecular recognition at a protein–small molecule or protein-protein interface has resulted in synthetic enzymes, kinases, transcription factors, and receptors with controllable biological functions, but here we apply this concept to protein interactions with cell surface receptors to control signaling specificity and downstream cellular functions (*21*–*28*). Orthogonal IL-2/IL-2R pairs may be useful not only as a research tool but in the clinic to specifically enrich transduced T cells that express a target gene of interest, such as a CAR or engineered TCR, when coupled with expression of the *ortho*IL-2Rβ. Our approach, and variations of this orthogonalization strategy, may be applicable to other cytokines, growth factors, hormones, and ligand-receptor interactions to decipher and manipulate otherwise complex biological systems.

REFERENCES AND NOTES

1. M. Kalos, C. H. June, *Immunity* **39**, 49–60 (2013).
2. S. A. Rosenberg, N. P. Restifo, *Science* **348**, 62–68 (2015).
3. C. Yee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16168–16173 (2002).
4. R. Andersen *et al.*, *Clin. Cancer Res.* **22**, 3734–3745 (2016).
5. S. A. Rosenberg, *J. Immunol.* **192**, 5451–5458 (2014).
6. X. Wang, M. Rickert, K. C. Garcia, *Science* **310**, 1159–1163 (2005).
7. A. M. Levin *et al.*, *Nature* **484**, 529–533 (2012).

8. A. B. Shanafelt *et al*., *Nat. Biotechnol.* **18**, 1197–1202 (2000).
9. M. Rickert, M. J. Boulanger, N. Goriatcheva, K. C. Garcia, *J. Mol. Biol.* **339**, 1115–1128 (2004).
10. E. Roessler *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3344–3347 (1994).
11. R. Moriggl *et al*., *Immunity* **10**, 249–259 (1999).
12. T. R. Malek, *Annu. Rev. Immunol.* **26**, 453–479 (2008).
13. D. Klatzmann, A. K. Abbas, *Nat. Rev. Immunol.* **15**, 283–294 (2015).
14. S. Amu, I. Gjertsson, A. Tarkowski, M. Brisslert, *Scand. J. Immunol.* **64**, 482–492 (2006).
15. O. Boyman, J. Sprent, *Nat. Rev. Immunol.* **12**, 180–190 (2012).
16. L. Gattinoni *et al*., *J. Exp. Med.* **202**, 907–912 (2005).
17. J. T. Sockolosky, F. C. Szoka, *Adv. Drug Deliv. Rev.* **91**, 109–124 (2015).
18. E. F. F. Zhu *et al*., *Cancer Cell* **27**, 489–501 (2015).
19. W. W. Overwijk *et al*., *J. Exp. Med.* **198**, 569–580 (2003).
20. K. Sakuishi *et al*., *J. Exp. Med.* **207**, 2187–2194 (2010).
21. M. G. J. Baud *et al*., *Science* **346**, 638–641 (2014).
22. S. Atwell, M. Ultsch, A. M. De Vos, J. A. Wells, *Science* **278**, 1125–1128 (1997).
23. D. M. Spencer, T. J. Wandless, S. L. Schreiber, G. R. Crabtree, *Science* **262**, 1019–1024 (1993).
24. D. J. Mandell, T. Kortemme, *Nat. Chem. Biol.* **5**, 797–807 (2009).
25. G. T. Kapp *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5277–5282 (2012).
26. J. S. Park *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 5896–5901 (2014).
27. S. M. Lewis *et al*., *Nat. Biotechnol.* **32**, 191–198 (2014).
28. C. Y. Wu, K. T. Roybal, E. M. Puchner, J. Onuffer, W. A. Lim, *Science* **350**, aab4077 (2015).

ACKNOWLEDGMENTS

We thank M. McCracken for expertise in T cell immunotherapy; N. Saligrama, A. Cravens, M. Hollander, F. Zhao, Y. Rosenberg-Hasson, and the Stanford Human Immune Monitoring Core (HIMC) for technical assistance; and R. Fernandes for helpful discussion. The data presented in this paper are tabulated in the main text and supplementary materials. Supported by NIH grants R37 AI051321 and HHMI (K.C.G.); a 2016 Stanford Cancer Institute translational research grant (J.T.S., A.R., and K.C.G.); NIH grant R35 CA197633, the Ressler Family Fund, and the Parker Institute for Cancer Immunotherapy (G.P. and A.R.); the Sean N. Parker Autoimmunity Research Laboratory (J.A.B.); and fellowship support from Stanford Molecular and Cellular Immunobiology NIH training grant 5T32 AI072905 and a PhRMA Foundation Translational Medicine and Therapeutics postdoctoral award (J.T.S.). J.A.B. and A.R. are members of the Parker Institute for Cancer Immunotherapy. K.C.G., J.T.S., I.C.K., and D.B. are inventors on patent applications 62/217,364 and 62/375,089 submitted by Stanford University that cover the use of orthogonal cytokine-receptor pairs for use in cellular immunotherapy.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1037/suppl/DC1
Materials and Methods
Figs. S1 to S18
References (*29*–*35*)

30 October 2017; accepted 11 January 2018
10.1126/science.aar3246

PROTEIN DESIGN

Accurate computational design of multipass transmembrane proteins

Peilong Lu,[1,2] Duyoung Min,[3] Frank DiMaio,[1,2] Kathy Y. Wei,[1,2]* Michael D. Vahey,[4] Scott E. Boyken,[1,2] Zibo Chen,[1,2] Jorge A. Fallas,[1,2] George Ueda,[1,2] William Sheffler,[1,2] Vikram Khipple Mulligan,[1,2] Wenqing Xu,[5] James U. Bowie,[3] David Baker[1,2,6]†

The computational design of transmembrane proteins with more than one membrane-spanning region remains a major challenge. We report the design of transmembrane monomers, homodimers, trimers, and tetramers with 76 to 215 residue subunits containing two to four membrane-spanning regions and up to 860 total residues that adopt the target oligomerization state in detergent solution. The designed proteins localize to the plasma membrane in bacteria and in mammalian cells, and magnetic tweezer unfolding experiments in the membrane indicate that they are very stable. Crystal structures of the designed dimer and tetramer—a rocket-shaped structure with a wide cytoplasmic base that funnels into eight transmembrane helices—are very close to the design models. Our results pave the way for the design of multispan membrane proteins with new functions.

In recent years, it has become possible to de novo design, with high accuracy, soluble protein structures ranging from short constrained peptides to megadalton protein cages (*1*). There have also been advances in membrane protein design, as illustrated by an elegant zinc-transporting transmembrane peptide tetramer named Rocker (*2*) and an engineered ion-conducting oligomer based on the C-terminal transmembrane segment (TMs) of the *Escherichia coli* polysaccharide transporter Wza (*3*). Both are single membrane–spanning synthesized peptides with fewer than 36 residues. It has also been possible to design and confirm the transmembrane topology of multipass membrane proteins by using simple sequence hydrophobicity and charge-based models (*4*), but the extent to which the transmembrane helices pack with each other is not clear. Design of structurally defined multipass membrane proteins has remained a major challenge because of the difficulty in specifying structure within the membrane and in experimentally determining membrane protein structures generally; crystal structures of the full designed oligomeric states of Rocker- and the Wza-derived channel have not yet been determined, and to date, there are no crystal structures of de novo–designed multipass membrane proteins.

A major challenge for membrane protein design stems from the similarity of the membrane environment to protein hydrophobic cores. In the design of soluble proteins, the secondary structure and overall topology can be specified by the pattern of hydrophobic and hydrophilic residues, with the former inside the protein and the latter outside, facing solvent. This core design principle cannot be used for membrane proteins because the apolar environment of the hydrocarbon core of the lipid bilayer requires that outward-facing residues in the membrane also be nonpolar. Buried hydrogen bonds between polar side chains have been demonstrated to play an important role in the association of helical peptides within the membrane, overcoming the degeneracy in the nonpolar interactions (*5*–*7*).

We reasoned that a recently developed method for designing buried hydrogen bond networks (*8*) could allow specification of the packing interactions of transmembrane helices in multipass transmembrane proteins. We first explored the design of helical transmembrane proteins with four TMs—dimers of 76- to 104-residue hairpins or a single chain design of 156 residues—with hydrophobic spanning regions ranging from 21 to 35 Å (Figs. 1A and 2A), repurposing the Ser- and Gln-containing hydrogen bond networks in a designed soluble four-helix dimer with C2 symmetry [2L4HC2_23; Protein Data Bank (PDB) ID 5J0K] (*8*) to provide structural specificity. Four-helix bundles of different lengths with backbone geometries capable of hosting these networks were produced by using parametric generating equations (*9*), residues comprising the hydrogen bond networks and neighboring packing residues were introduced, and the remainder of the sequence was optimized by using Rosetta Monte Carlo (*10*) design calculations to obtain low-energy sequences. Connecting loops between the helices were built with Rosetta. To specify the orientation of the designs (*11*) in the membrane when expressed in cells, at the designed lipid-water boundary on the extracellular/periplasmic side, we incorporated a ring of amphipathic aromatic residues and, at the lipid-water boundary on the cytoplasmic side, a ring of positively charged residues (Figs. 1A and 2A). Between these two rings, the surface residues are exposed to the hydrophobic membrane environment; these positions in Rosetta sequence design calculations were restricted to hydrophobic amino acids (supplementary materials). Consistent with the design, TMHMM predicts that the dimer designs contain 2 TMs and the single-chain design (scTMHC2) contains 4 TMs (fig. S1). On average, for each residue ~68% of the side-chain surface area is buried in the design models, which could provide substantial van der Waals stabilization (*12*).

Synthetic genes encoding the designs were obtained and the proteins expressed in *E. coli* and mammalian cells. The dimer design with the shortest hydrophobic span (15 residues; TMHC2_S) was poorly behaved in both *E. coli* and mammalian cells, but the dimer designs with longer spans—TMHC2, TMHC2_E, and TMHC2_L—localized to the cell membrane when expressed in human embryonic kidney (HEK) 293T cells (Fig. 1B) and in *E. coli*. The designed proteins were purified by extracting the *E. coli* membrane fraction with detergent, followed by nickel–nitrilotriacetic acid (NTA) chromatography and size exclusion chromatography (SEC) with a yield of ~2 mg/L (fig. S2, A and B). The designed proteins TMHC2, TMHC2_E, and TMHC2_L eluted as single peaks in SEC, and in analytical ultracentrifugation (AUC) experiments in detergent solution, the proteins sedimented as dimers, which is consistent with the design models (Fig. 1C and fig. S3). For the single-chain scTMHC2, the major species in SEC was the monomer, with a small side peak that was readily removed by purification (fig. S2B). Circular dichroism (CD) measurements showed that the designs were α-helical and highly thermal stable; the CD spectra at 95°C were similar to those at 25°C (Figs. 1D and 2B). TOXCAT-β–lactamase (TβL) assays (*13*), which couple *E. coli* survival to oligomerization and proper orientation of fused antibiotic resistance markers on the N and C termini, suggest that the N and C termini of TMHC2 are in the cytoplasm, as in the design models (fig. S4).

We more quantitatively characterized the folding stability of scTMHC2 using single-molecule forced unfolding experiments (Fig. 2) (*14*, *15*). The designed protein reconstituted in a bicelle was covalently attached to a magnetic bead and a glass surface through its N and C termini (Fig. 2A and fig. S5). The distance between the bead and the surface was determined as a function of the applied mechanical tension. In unfolding experiments with the force slowly increasing (~0.5 pN/s), unfolding transitions were observed at ~18 pN and, upon force deramping, refolding transitions were observed

[1]Department of Biochemistry, University of Washington, Seattle, WA 98195, USA. [2]Institute for Protein Design, University of Washington, Seattle, WA 98195, USA. [3]Department of Chemistry and Biochemistry, UCLA-DOE Institute, Molecular Biology Institute, University of California, Los Angeles (UCLA), CA 90095, USA. [4]Department of Bioengineering and Biophysics Group, University of California, Berkeley, Berkeley, CA 94720, USA. [5]Department of Biological Structure, University of Washington, Seattle, WA 98195, USA. [6]Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA.
*Present address: Department of Bioengineering, University of California, Berkeley, CA 94720, USA.
†Corresponding author. Email: dabaker@u.washington.edu

at ~9 pN (80.1% of the recorded unfolding traces had one-step unfolding transitions, and 84.6% of the refolding transitions had two steps) (Fig. 2C and figs. S6 and S7). Consistent with the internal symmetry of the single-chain design (Fig. 2A and fig. S5), the two refolding step sizes were very similar (fig. S8). This unfolding and refolding asymmetry is consistent with a three-state free-energy landscape: the native state (*N*), an intermediate state containing only one hairpin (*I*), and an unfolded state (*U*) (fig. S9). During unfolding at high force, only the barrier between the native and intermediate states is observed, whereas at the lower forces at which refolding occurs, both energy barriers become

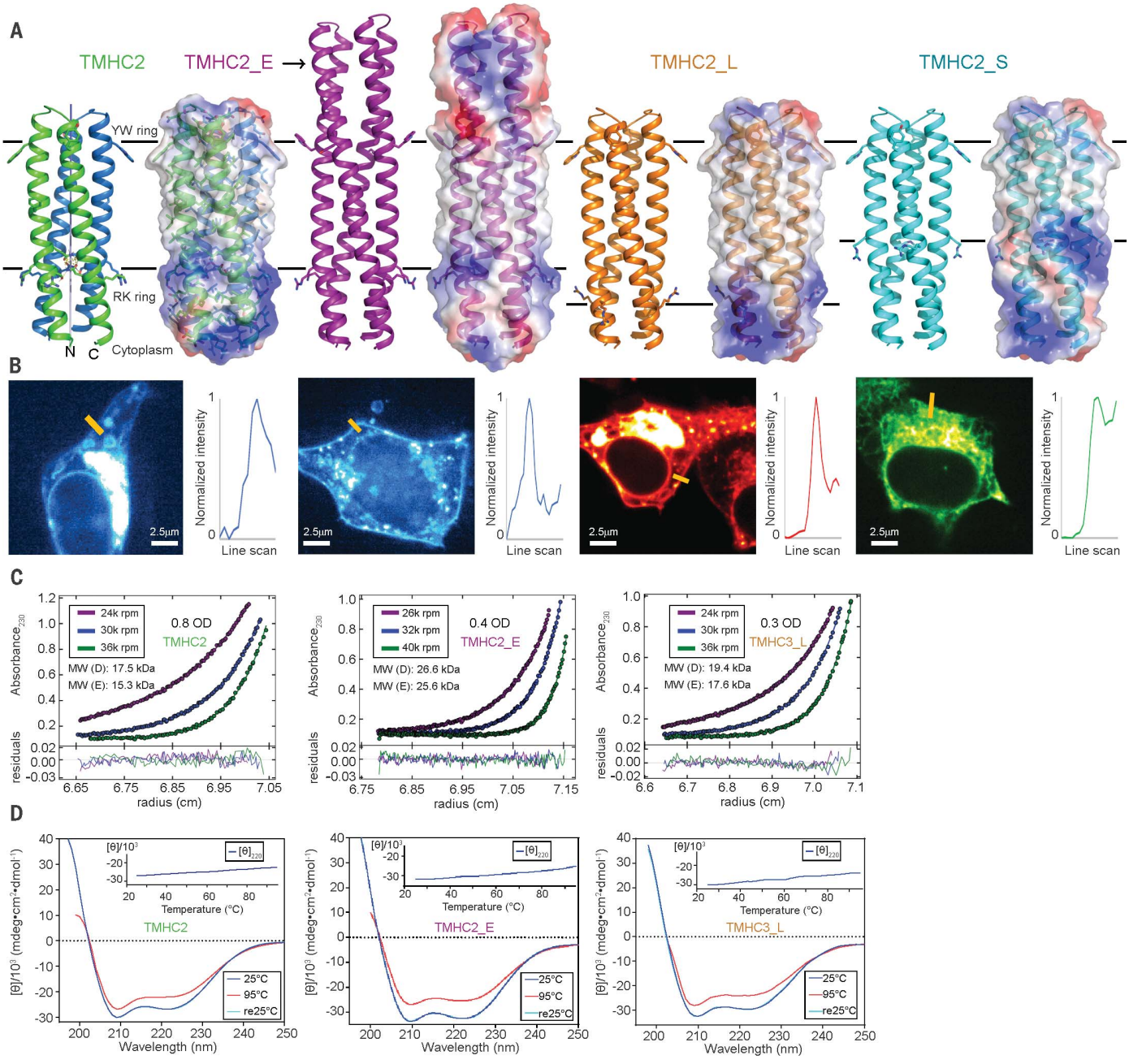


Fig. 1. Design and characterization of proteins with four transmembrane helices. (**A** and **B**) From left to right, designs and data for TMHC2 (transmembrane hairpin C2), TMHC2_E (elongated), TMHC2_L (long span), and TMHC2_S (short span). (A) Design models with intra- and extramembrane regions with different lengths. Horizontal lines demarcate the hydrophobic membrane regions. Ribbon diagrams are at left, electrostatic surfaces are at right, and the neutral transmembrane regions are in gray. (B) Confocal microscopy images for HEK293T cells transfected with TMHC2 fused to mTagBFP, TMHC2_E fused to mTagBFP, TMHC2_L fused to mCherry, and TMHC2_S fused to enhanced green fluorescent protein. Line scans (yellow lines) across the membranes show substantial increase in fluorescence across the plasma membranes for TMHC2, TMHC2_E, and TMHC2_L, but less substantial increase for TMHC2_S. (**C**) Representative AUC sedimentation-equilibrium curves at three different rotor speeds. Each data set is globally well fit as a single ideal species in solution corresponding to the dimer molecular weight. "MW (D)" and "MW (E)" indicate the molecular weight of the oligomer design and that determined from experiment, respectively. (**D**) CD spectra and (inset) temperature melt. No apparent unfolding transitions are observed up to 95°C.

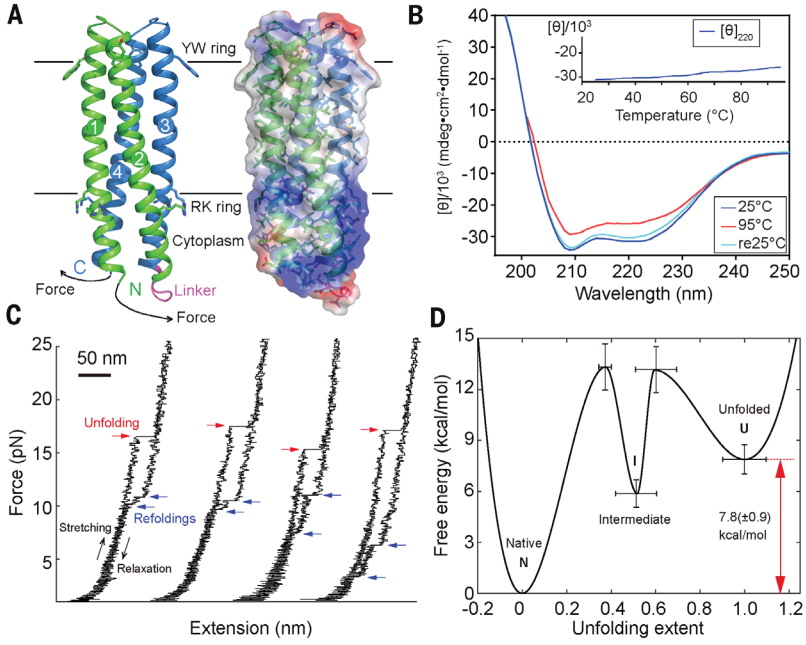


Fig. 2. Folding stability of the 156-residue single-chain TMHC2 (scTMHC2) design with four transmembrane helices. (**A**) Design model (left) and electrostatic surface (right) of scTMHC2. N- and C-terminal helical hairpins are colored green and blue, respectively. Numbers indicate the order of the four TMs in the sequence. The linker connecting the two hairpins is colored magenta. Single-molecule forced unfolding experiments were conducted by applying mechanical tension to the N and C termini of a single scTMHC2 (fig. S5). (**B**) CD spectra of scTMHC2 at different temperatures. No unfolding transition is observed up to 95°C. (**C**) Single-molecule force-extension traces of scTMHC2. The unfolding and refolding transitions are denoted with red and blue arrows. (**D**) Folding energy landscape obtained from the single-molecule experiments. *N*, *I*, and *U* indicate the native, intermediate, and unfolded state, respectively.

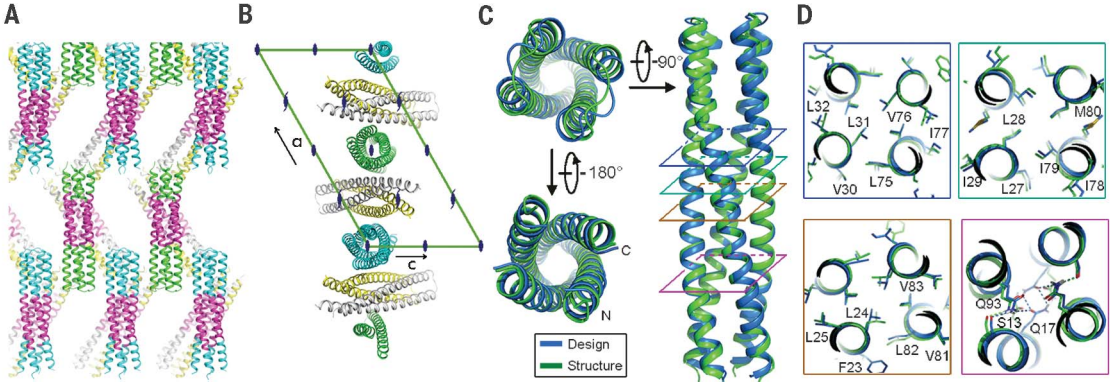


Fig. 3. Crystal structure of the designed transmembrane dimer TMHC2_E. (**A** and **B**) Crystal lattice packing. (A) The extended soluble region mediates a large portion of the crystal lattice packing. The four helical hairpins in the asymmetric unit are colored green, gray, yellow, and blue, respectively. The TMs, in magenta, forms layers in the crystal separating the soluble regions. (B) The C2 axis of the design aligns with the crystallographic twofold. Two monomers (gray and yellow) are paired in a dimer, whereas the other two (green and blue) form two C2 dimers with two crystallographic adjacent monomers. The space group diagram (C121) is shown in the background. (**C**) Superposition of the TMHC2_E crystal structure and design model (RMSD = 0.7 Å over the core Cα atoms). (**D**) The side-chain packing arrangements at layers [(C), colored squares] at different depths in the membrane are almost identical to the design model.

prominent (fig. S9). The transition rates between the folded, intermediate, and unfolded states were determined by using the Bell model (*16*), yielding the relative free energies of the states and the associated barrier heights (Fig. 2D and fig. S10) (*14*). The overall thermodynamic stability of scTMHC2 is 7.8(±0.9) kcal/mol on a per transmembrane helix basis, which is more stable than the naturally occurring helical membrane proteins studied thus far [folding free energy per helix for scTMHC2 is 2.0(±0.2) kcal/(mol·helix) compared with 0.7 to 0.9 kcal/(mol·helix) for GlpG (*14*, *17*) and 1.6 to 1.8 kcal/(mol·helix) for bacteriorhodopsin (*18*); error estimates in parentheses are propagated from the standard errors of the kinetics measurements].

We carried out crystal screens in different detergents for each of the designs and obtained crystals of the design with the most extensive cytoplasmic region, TMHC2_E, in *n*-nonyl-β-D-glucopyranoside (NG). The crystals diffracted to 2.95-Å resolution, and we solved the structure by means of molecular replacement with the design model. As anticipated, the extended soluble region mediates the crystal lattice packing; there are large solvent channels around the designed TMs likely because of the surrounding disordered detergent molecules (Fig. 3A). Each asymmetric unit contains four helical hairpins: Two are paired in a dimer, whereas the other two form two C2 dimers through crystallographic symmetry with two monomers in adjacent asymmetric units. The C2 axis in the design is perfectly aligned with the crystallographic twofold (Fig. 3B). The conformations of the dimers in the three biological units are nearly identical, with very small differences due to crystal packing [Cα root-mean-square deviations (RMSDs), 0.60 to 0.84 Å] (fig. S11). Both the overall structure and the core side-chain packing are almost identical in the crystal structure and the design model, with a Cα RMSD of 0.7 Å over the core residues (Fig. 3C). Two of the three buried hydrogen bonding residues within the membrane have conformations that almost exactly match the design model (S13 and Q93), but Q17 adopts a different rotamer, with the side-chain nitrogen donating a hydrogen bond to the main-chain carbonyl oxygen (Fig. 3D).

We used a similar approach to design a transmembrane trimer with six membrane-spanning helices (TMHC3) based on the 5L6HC3_1 scaffold (PDB ID 5IZS) (*8*). Guided by the results with the C2 designs, we chose a hydrophobic span of ~30 Å (20 residues) (Fig. 4A). The design was expressed in *E. coli* and purified to

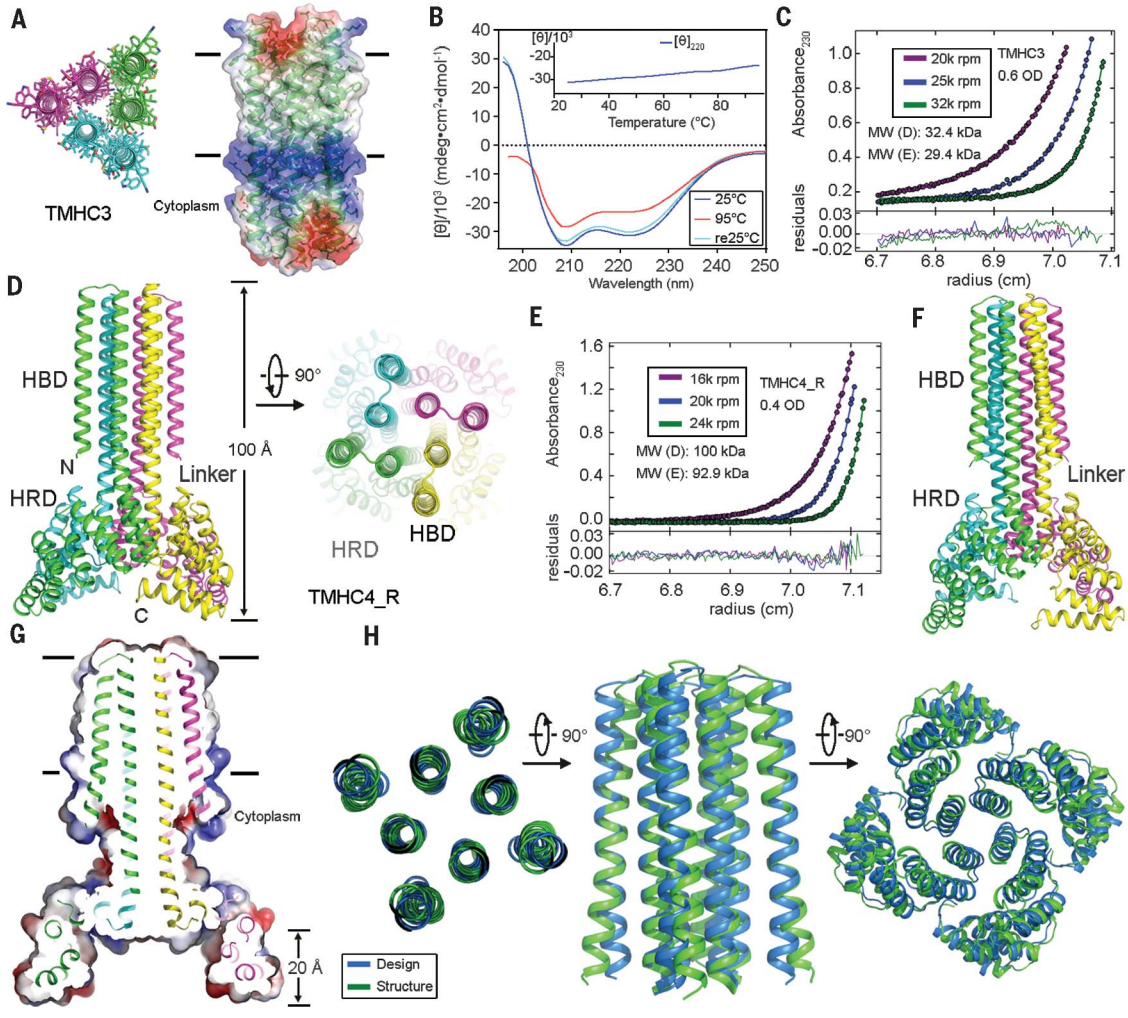


Fig. 4. Stability and structural characterization of designs with six and eight membrane-spanning helices. (**A**) Model of designed transmembrane trimer TMHC3 with six transmembrane helices. Stick representation from periplasmic side (left) and lateral surface view (right) are shown. (**B**) CD characterization of TMHC3. The design is stable up to 95°C. (**C**) Representative AUC sedimentation-equilibrium curves at three different rotor speeds for TMHC3. The data fit to a single ideal species in solution with molecular weight close to that of the designed trimer. (**D**) Model of designed transmembrane tetramer TMHC4_R with eight transmembrane helices. The four protomers are colored green, yellow, magenta, and blue, respectively. (**E**) AUC sedimentation-equilibrium curves at three different rotor speeds for TMHC4_R fit well to a single species, with a measured molecular weight of ~94 kDa. (**F**) Crystal structure of TMHC4_R. The overall tetramer structure is very similar to the design model, with a helical bundle body and helical repeat fins. The outer helices of the transmembrane hairpins tilt off the axis by ~10°. (**G**) Cross section through the TMHC4_R crystal structure and electrostatic surface. The HRD forms a bowl at the base of the overall structure with a depth of ~20 Å. The transmembrane region is indicated in lines. (**H**) Three views of the backbone superposition of TMHC4_R crystal structure and design model.

homogeneity, eluting on a gel filtration column as a single homogeneous species (fig. S2C). CD measurements showed that TMHC3 was highly thermostable, with the α-helical structure preserved at 95°C (Fig. 4B). AUC experiments showed that TMHC3 is a trimer in detergent solution, which is consistent with the design model (Fig. 4C and fig. S12A).

To explore our capability to design membrane proteins with more complex topologies, we designed a C4 tetramer with a two-ring, helical membrane-spanning region composed of eight TMs and an extended bowl-shaped cytoplasmic domain formed by repeating structures emanating away from the symmetry axis (Fig. 4D). The design has an overall rocket shape, with a height of ~100 Å, and can be divided into three regions: the helical bundle domain (HBD), the helical repeat domain (HRD), and the helical linker between the two. The central HBD was derived from the soluble design 5L8HC4_6 (*8*), and the bowl was derived from a designed helical repeat protein homo-oligomer (tpr1C4_2) (*19*). Helical linkers were built by using RosettaRemodel (*20*); a nine-residue junction was found to yield the correct helical register (fig. S13). After Rosetta sequence design calculations, a gene encoding the lowest energy design, TMHC4_R, was synthesized. The protein was expressed in *E. coli* and purified by using nickel affinity and gel filtration chromatography; the final yield was ~3 mg/L, and the purified protein chromatographed as a monodisperse peak in SEC (fig. S2C). CD experiments showed that the design was α-helical and thermostable up to 95°C (fig. S12B). AUC measurements showed that TMHC4_R is a tetramer in detergent solution, which is consistent with the design model (Fig. 4E and fig. S12C). After a systematic effort to screen detergents for crystallization, we obtained crystals in a combination of *n*-decyl-β-D-maltopyranoside (DM) and NG in the P4 space group that diffracted to 3.9-Å resolution. We solved the crystal structure by means of molecular replacement using the design model (R_{work}/R_{free} = 0.29/0.32, with unambiguous electron density) (table S1 and fig. S14). The crystal lattice packing is primarily between the extended cytoplasmic domains; there may be minor detergent-mediated interactions between the transmembrane and helical repeat (HR) domains as well (fig. S15).

Although the resolution is insufficient for evaluating the details of the side-chain packing, it does allow backbone-level comparisons. There are four TMHC4_R monomers in one asymmetric unit, with nearly identical structures (Cα RMSDs between 0.2 and 0.6 Å) (fig. S16A). The Cα RMSDs between the structure and design model are 1.2 to 1.8 Å for the monomer transmembrane helices, 0.3 to 0.4 Å for the linkers, 1.1 to 1.5 Å for the HR domains, and 3.3 to 3.6 Å for the overall structure (fig.

S16B). As in the case of the C2 design, the C4 symmetry axis of the design coincides with the crystallographic axes of the crystal lattice (fig. S16C). The four tetramer structures on the crystal C4 axes have overall structures very similar to each other and to the design model (Fig. 4, F and G, and fig. S16A); the tetrameric transmembrane domain, HR domain, and overall tetramer structure have Cα RMSDs to the design model of 1.3 to 1.5 Å, 3.3 to 3.8 Å, and 3.3 to 3.8 Å, respectively (Fig. 4H and fig. S16D, left). The deviation in the HR domain may result from crystal packing interactions between the termini; the Cα RMSDs over the first 162 residues are 2.2 to 2.3 Å (fig. S16D, right). The main deviation from the design model is a tilting of the outer helices of transmembrane hairpins from the axis by ~10° (Fig. 4, F and G).

The agreement between the crystal structures of TMHC2_E and TMHC4_R with the design models demonstrates that transmembrane homo-oligomers containing multiple membrane-spanning regions and extensive extracellular domains can now be accurately designed. For future work, the general approach of first designing and characterizing hydrogen bond network–containing soluble versions of the desired transmembrane structures, and then converting to integral membrane proteins by redesigning the membrane-exposed residues, could be quite robust. Single-molecule forced unfolding and thermal denaturation experiments show that the designed proteins are highly stable. Although the designs lack the classic small residue packing in the core that is thought to be an important driver of membrane protein folding (*21*–*24*), like natural membrane proteins they bury more surface area than do typical soluble proteins, maximizing van der Waals packing contributions (*12*). The range of the design features—variable transmembrane and extracellular helix lengths and superhelical twists, extensive soluble domains, and diverse oligomeric states—are substantial steps toward the complexity of natural transmembrane proteins with multiple membrane-spanning regions and extra-membrane domains that play important roles in ligand/substrate recognition and structure stabilization, such as in the adenosine 5′-triphosphate–binding cassette transporters, ion channels, ryanodine receptor, and γ-secretase (*25*, *26*). The capability to accurately design complex multipass transmembrane proteins that can be expressed in cells opens the door to the design of a new generation of multipass membrane protein structures and functions.

REFERENCES AND NOTES

1. P.-S. Huang, S. E. Boyken, D. Baker, *Nature* **537**, 320–327 (2016).
2. N. H. Joh *et al.*, *Science* **346**, 1520–1524 (2014).
3. K. R. Mahendran *et al.*, *Nat. Chem.* **9**, 411–419 (2017).
4. P. Whitley, I. Nilsson, G. von Heijne, *Nat. Struct. Biol.* **1**, 858–862 (1994).
5. C. Choma, H. Gratkowski, J. D. Lear, W. F. DeGrado, *Nat. Struct. Biol.* **7**, 161–166 (2000).
6. F. X. Zhou, M. J. Cocco, W. P. Russ, A. T. Brunger, D. M. Engelman, *Nat. Struct. Biol.* **7**, 154–160 (2000).
7. C. D. Tatko, V. Nanda, J. D. Lear, W. F. Degrado, *J. Am. Chem. Soc.* **128**, 4170–4171 (2006).
8. S. E. Boyken *et al.*, *Science* **352**, 680–687 (2016).
9. P.-S. Huang *et al.*, *Science* **346**, 481–485 (2014).
10. C. A. Rohl, C. E. M. Strauss, K. M. S. Misura, D. Baker, in *Methods in Enzymology* (2004), pp. 66–93.
11. G. von Heijne, *Nat. Rev. Mol. Cell Biol.* **7**, 909–918 (2006).
12. A. Oberai, N. H. Joh, F. K. Pettit, J. U. Bowie, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 17747–17750 (2009).
13. A. Elazar *et al.*, *eLife* **5**, e12125 (2016).
14. D. Min, R. E. Jefferson, J. U. Bowie, T.-Y. Yoon, *Nat. Chem. Biol.* **11**, 981–987 (2015).
15. D. Min, M. A. Arbing, R. E. Jefferson, J. U. Bowie, *Protein Sci.* **25**, 1535–1544 (2016).
16. G. I. Bell, *Science* **200**, 618–627 (1978).
17. R. Guo *et al.*, *Nat. Chem. Biol.* **12**, 353–360 (2016).
18. Y.-C. Chang, J. U. Bowie, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 219–224 (2014).
19. J. A. Fallas *et al.*, *Nat. Chem.* **9**, 353–360 (2017).
20. P.-S. Huang *et al.*, *PLOS ONE* **6**, e24109 (2011).
21. S.-Q. Zhang *et al.*, *Structure* **23**, 527–541 (2015).
22. R. F. S. Walters, W. F. DeGrado, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13658–13663 (2006).
23. X. Feng, P. Barth, *Nat. Chem. Biol.* **12**, 167–173 (2016).
24. A. D. Meruelo, S. K. Han, S. Kim, J. U. Bowie, *Protein Sci.* **21**, 1746–1753 (2012).
25. Y. Shi, *Cell* **159**, 995–1014 (2014).
26. R. Fernandez-Leiro, S. H. W. Scheres, *Nature* **537**, 339–346 (2016).

ACKNOWLEDGMENTS

We thank J. Sumida for AUC support; A. Kang for crystallization support; D. Ma and Z. Wang for crystallography support; and the staff at the Advanced Light Source and P. Huang, Y. Hsia, A. Ford, L. Stewart, C. Xu, and many other members of the Baker laboratory for helpful discussions. This work was facilitated by the Hyak supercomputer at the University of Washington. **Funding:** This work was supported by the Howard Hughes Medical Institute (D.B.) and the National Institutes of Health (grant R01GM063919 to J.U.B.). P.L. was supported by the Raymond and Beverly Sackler fellowship. D.M. was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (grant NRF-2016R1A6A3A03007871). **Author contributions:** P.L. and D.B. designed the research, and P.L., D.M., F.D., K.Y.W., J.U.B., and D.B. wrote the manuscript. P.L. and D.B. carried out design calculations and developed the membrane protein design method. P.L. purified and characterized the designed proteins. D.M. and J.U.B. designed, performed, and analyzed single-molecule forced unfolding experiments. K.Y.W. and M.D.V. performed mammalian cell localization experiment. P.L. crystallized the designed proteins. P.L. collected and analyzed crystallographic data with help from W.X.; F.D. solved structures with help from P.L.; and S.E.B., C.Z., J.A.F., G.U., and W.S. contributed the soluble scaffolds. V.K.M. wrote the amino acid composition–based energy term. All authors discussed results and commented on the manuscript. **Competing interests:** D.B., P.L., S.E.B., C.Z., J.A.F., G.U., and W.S. are inventors on a U.S. provisional patent application submitted by the University of Washington that covers computational design of multipass transmembrane proteins described in this paper. **Data and materials availability:** Coordinates and structure files have been deposited to PDB with accession codes 6B87 (TMHC2_E) and 6B85 (TMHC4_R). All other data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1042/suppl/DC1
Materials and Methods
Figs. S1 to S16
Table S1
References (*27*–*39*)

10 October 2017; accepted 11 January 2018
10.1126/science.aaq1739

NEUROPHYSIOLOGY

An evolutionarily conserved gene family encodes proton-selective ion channels

Yu-Hsiang Tu,[1] Alexander J. Cooper,[1]*† Bochuan Teng,[1]* Rui B. Chang,[1]*‡ Daniel J. Artiga,[1] Heather N. Turner,[1] Eric M. Mulhall,[1]§ Wenlei Ye,[1]‖ Andrew D. Smith,[2] Emily R. Liman[1,3]#

Ion channels form the basis for cellular electrical signaling. Despite the scores of genetically identified ion channels selective for other monatomic ions, only one type of proton-selective ion channel has been found in eukaryotic cells. By comparative transcriptome analysis of mouse taste receptor cells, we identified Otopetrin1 (OTOP1), a protein required for development of gravity-sensing otoconia in the vestibular system, as forming a proton-selective ion channel. We found that murine OTOP1 is enriched in acid-detecting taste receptor cells and is required for their zinc-sensitive proton conductance. Two related murine genes, *Otop2* and *Otop3*, and a *Drosophila* ortholog also encode proton channels. Evolutionary conservation of the gene family and its widespread tissue distribution suggest a broad role for proton channels in physiology and pathophysiology.

Ion channels include a large and diverse group of membrane proteins that rapidly, and with great selectivity, move ions across the cell membrane, performing crucial roles in cell signaling and homeostasis (*1*). Ion channels selective for each of the physiologically relevant ions, Na^+, K^+, Ca^{2+}, and Cl^-, have been described at the molecular and structural levels (*2*, *3*), but only a few types of proton-selective ion channels (proton channels) have been described (*4*). One is the 96–amino acid M2 protein of influenza A, which conducts protons into the virion interior, an essential step in the replication of the virus (*5*). The only proton-selective ion channel identified in eukaryotes is the voltage-gated Hv1 (*6*–*8*), which is present in immune cells, where it extrudes protons into the phagosome to inactivate infectious agents (*9*). Functional evidence indicates that ion channels that selectively transport protons into eukaryotic cells must also exist. For example, in acid-sensing taste receptor cells (TRCs), an inward-conducting Zn^{2+}-sensitive proton current that is biophysically distinct from currents carried by Hv1 has been described (*10*, *11*).

To identify candidates encoding such a proton channel, we compared the transcriptome of mouse TRCs positive for the inward-conducting Zn^{2+}-sensitive proton current (PKD2L1 cells) with that of TRCs that lack the current (TRPM5 cells; Fig. 1A). We selected genes that were enriched in PKD2L1 cells and that encoded poorly characterized or uncharacterized transmembrane proteins (Fig. 1A and table S1) (see methods). We expressed the candidates in human embryonic kidney 293 (HEK-293) cells or *Xenopus* oocytes and measured ionic currents in response to lowering the extracellular pH (pH_o) in the absence of extracellular Na^+. Of the 41 cDNAs tested, only Otopetrin1 (*Otop1*), which encodes a protein (OTOP1) with 12 predicted transmembrane domains (*12*), generated large Zn^{2+}-sensitive inward currents in response to extracellular acidification (Fig. 1B).

We characterized functional properties of OTOP1 expressed in *Xenopus* oocytes. Unless otherwise noted, the extracellular solution used in recordings

[1]Department of Biological Sciences, Section of Neurobiology, University of Southern California, Los Angeles, CA 90089, USA. [2]Department of Biological Sciences, Section of Molecular and Computational Biology, University of Southern California, Los Angeles, CA 90089, USA. [3]Bridge Institute, University of Southern California, Los Angeles, CA 90089, USA.
*These authors contributed equally to this work. †Present address: Zilkha Neurogenetic Institute, Department of Cell and Neurobiology, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA. ‡Present address: Department of Neuroscience and Department of Cellular and Molecular Physiology, Yale University School of Medicine, New Haven, CT 06520, USA. §Present address: Department of Neurobiology, Harvard Medical School, Boston, MA 20115, USA. ||Present address: Department of Physiology, University of California, San Francisco, CA 94158, USA.
#Corresponding author. Email: liman@usc.edu

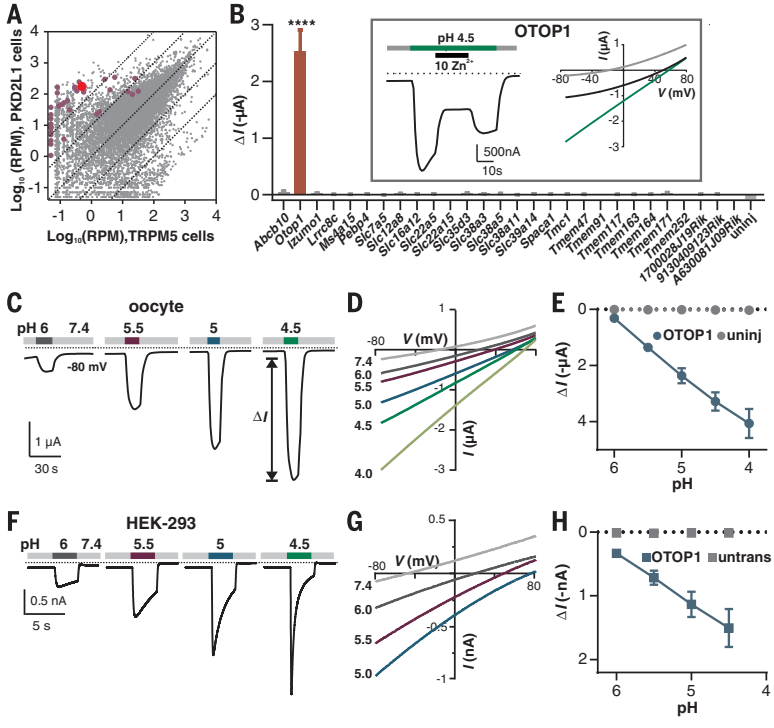
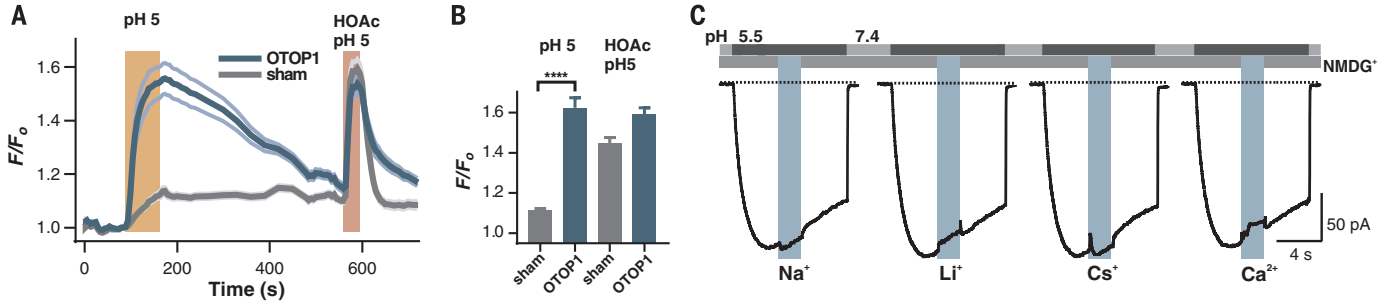


Fig. 1. Expression analysis of taste-cell–enriched genes identifies OTOP1 as a previously unknown proton channel. (**A**) Transcriptome profiling of PKD2L1 and TRPM5 taste receptor cells (each data point represents the average of five replicates). Genes tested by electrophysiology are highlighted in magenta or red (*Otop1*). RPM, reads per million. (**B**) Magnitude of currents evoked in response to pH 4.5 Na^+-free solution in *Xenopus* oocytes expressing the genes indicated (V_m = −80 mV; data are mean ± SEM, n = 3 to 37 cells; for OTOP1, n = 5). ****P < 0.0001 compared to uninjected oocytes (n = 3). One-way analysis of variance with Bonferroni correction. (Inset) Currents evoked in an OTOP1-expressing oocyte in response to the acid stimulus at V_m = −80mV (left) and the current-voltage (I-V) relationship before application (gray), during acid application (green), and during Zn^{2+} application (black). (**C**) Current measured by two-electrode voltage clamp in a *Xenopus* oocyte expressing OTOP1 in response to Na^+-free extracellular solutions with pH_o as indicated (V_m = −80 mV). (**D**) I-V relation of the current in (A) from voltage ramps (1 V/s). (**E**) Evoked current (ΔI; mean ± SEM) as a function of pH in *Xenopus* oocytes expressing OTOP1 (blue circle; n = 4) and uninjected oocytes (gray circles; n = 4). (**F**) Currents measured by whole-cell patch clamp recording in a HEK-293 cell expressing OTOP1 in Na^+-free extracellular solutions (pH_i = 7.3, V_m = −80 mV). (**G**) I-V relation of currents in an OTOP1-expressing HEK-293 cell from experiments as in (G) with voltage ramps (1 V/s). (**H**) Evoked currents (ΔI; mean ± SEM) as a function of pH in HEK-293 cells expressing OTOP1 (blue squares; n = 5) and untransfected cells (gray squares; n = 3).



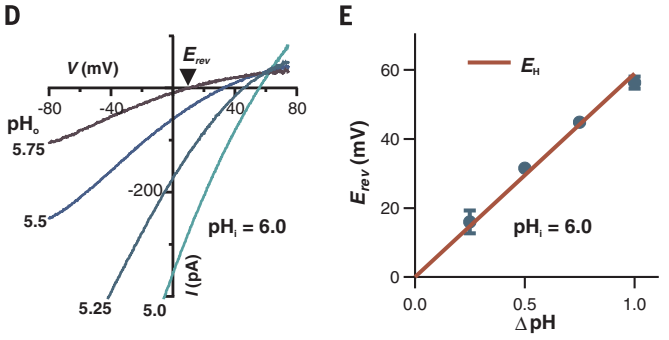


Fig. 2. Selectivity of OTOP1 for protons. (**A**) Fluorescence emission of the pH indicator pHrodo Red in HEK-293 cells expressing OTOP1 (n = 9) and sham-transfected cells (n = 11; mean ± SEM) in response to the stimuli indicated. Similar results were obtained in three replicates. (**B**) Average data (mean ± SEM; n = 28 or 29 cells) were analyzed by two-tailed t test. ****P < 0.0001. HOAc (acetic acid), which shuttles protons across membranes (*4*), served as a positive control. (**C**) OTOP1 currents in HEK-293 cells were evoked in response to a pH 5.5 solution with Na^+, Li^+, or Cs^+ (160 mM each) or Ca^{2+} (40 mM) replacing $NMDG^+$ in the extracellular solution as indicated (V_m = −80 mV). Percentage change in currents was 0.4 ± 0.7 (n = 8), 2.7 ± 0.7 (n = 8), 2.4 ± 0.5 (n = 8), and 3.6 ± 1.7 (n = 7) for each ion replacement, respectively. (**D**) Isolated OTOP1 currents in response to voltage ramps (1 V/s) at varying extracellular pH (pH_i = 6.0; Zn^{2+}-sensitive component is shown; see fig. S5 and methods). (**E**) E_{rev} as a function of ΔpH (pH_i − pH_o) from experiments as in (D). The red line is the equilibrium potential for H^+, E_H. The data were fit by linear regression with a slope of 53 mV/ΔpH and a y intercept of 3.6 mV (correlation coefficient R^2 = 0.99).

was Na^+-free [*N*-methyl-D-glucamine ($NMDG^+$)–based]. OTOP1 currents increased monotonically as pH_o was lowered (Fig. 1, C to E) and the reversal potential (E_{rev}) shifted toward more positive voltages (Fig. 1D and fig. S1A). The currents showed a small time-dependent change in amplitude in response to hyperpolarizing voltage steps, indicating that gating of OTOP1 is mildly voltage-sensitive (fig. S1, B and C).

OTOP1 also generated an ionic current in HEK-293 cells (Fig. 1, F to H). An N-terminal YFP (yellow fluorescent protein)–tagged protein confirmed the presence of OTOP1 at the cell surface (fig. S2A). Lowering pH_o elicited large inward currents in OTOP1-expressing cells and, as in oocytes, the current magnitude increased monotonically with pH_o (Fig. 1, F to H). OTOP1 currents in HEK-293 cells decayed within seconds, with faster kinetics observed in response to more acidic stimuli (Fig. 1F). The decay of the currents is likely to be due, in part, to a reduction in the driving force as protons accumulate in the cytosol. For a 15-μm-diameter cell (1767 fl volume), a H^+ current of 1000 pA flowing for 1 s will increase the total (bound + free) intracellular concentration of H^+ by ~6 mM (*4*). We confirmed that OTOP1 mediated flux of protons into the cell cytosol with the membrane-permeant pH indicator pHrodo Red. In *Otop1*-transfected cells, but not in mock-transfected cells, lowering extracellular pH from 7.4 to 5.0 caused a large increase in emission of pHrodo Red (Fig. 2, A and B), corresponding to a large change in intracellular pH (fig. S2, B and C).

Hv1 and M2 are highly selective for protons, present in high nanomolar concentrations, over other cations whose concentrations are a million times higher (*4*, *13*). To determine if OTOP1 is similarly proton-selective, we evoked Otop1 currents by lowering pH_o from 7.4 to 5.5 and measured the effect of exchanging $NMDG^+$ in the extracellular solution for equimolar concentrations of Na^+, Cs^+, or Li^+ or isosmotic concentrations of Ca^{2+} (Fig. 2C). In all cases, the observed change in current magnitude was less than 4%, indicating that OTOP1 is not appreciably permeable to these ions. Similar experiments showed that OTOP1 is not appreciably permeable to K^+ (fig. S3). To directly assess the selectivity of the channel for protons, we measured the potential at which the current reversed direction, the reversal potential (E_{rev}), as a function of the H^+ gradient (ΔpH = pH_i − pH_o). To study a predominantly OTOP1 current, we applied Zn^{2+} at a concentration that selectively and fully blocked the OTOP1 current in HEK-293 cells and focused on the Zn^{2+}-sensitive component of the current (figs. S4 and S5, A and B). We limited H^+ accumulation by setting pH_i at 6.0 and holding the membrane potential at E_{rev} (fig. S5A; see methods). Under these conditions, E_{rev} closely followed the Nernst prediction for an H^+-selective ion channel (Fig. 2, D and E). To determine the selectivity of OTOP1 for H^+ relative to Na^+ and Cl^-, we measured E_{rev} upon replacement of $NMDG^+$ by Na^+ and with high and low concentrations of Cl^- in the extracellular solution. In no case did we observe any change in E_{rev} (fig. S5C). Assuming a change in E_{rev} of less than 5 mV, which would have been detectable, we used the Goldman-Hodgkin-Katz equation to calculate the selectivity of OTOP1 for H^+ relative to Na^+ at greater than 2×10^5–fold and H^+ to Cl^- at greater than 1×10^5–fold (*13*).

OTOP1 is a member of the *otopetrin* family of proteins, which is evolutionarily conserved from nematodes to humans (*12*, *14*) (Fig. 3A). We confirmed that human OTOP1 (hOTOP1) forms a channel with properties similar to those of murine OTOP1 (fig. S6). Murine OTOP2 and OTOP3 share 30 to 34% amino acid identity with murine OTOP1 (fig. S7). Each shows a distinctive pattern of expression. *Otop1* is expressed in vestibular and taste cells, brown adipose tissue (*15*), heart, uterus, dorsal root ganglion, adrenal gland, mammary gland, and stimulated mast cells, whereas *Otop2* expression is highest in stomach, testis, and olfactory bulb, and *Otop3* is expressed in epidermis, small intestine, stomach, and retina [Fig. 3B; (*16*)]. When expressed in *Xenopus* oocytes, OTOP2 and OTOP3 both generated large currents in response to lowering pH_o in a Na^+-free solution. Compared with OTOP1 and OTOP3, OTOP2 currents behaved anomalously; currents saturated at ~pH 5, and E_{rev} shifted little over a range of pH 4 to 6 (Fig. 3, C to E, and fig. S8A). OTOP2 currents measured in HEK-293 cells had similar properties (fig. S9). Like OTOP1, OTOP3 showed evidence of selectivity for H^+; the magnitude of OTOP3 currents increased linearly as a function of pH_o over the entire pH range tested (Fig. 3, F to H), and E_{rev} shifted 46.3 mV/log[H^+], close to the value of 58 mV/log[H^+] expected for a proton-selective ion channel (fig. S8C). In response to hyperpolarizing voltage steps, OTOP2 and OTOP3 currents showed evidence of mild (OTOP3) or no (OTOP2) voltage dependence (fig. S8, B and D). When expressed in HEK-293 cells and assessed with microfluorimetry, both OTOP2 and OTOP3 conducted protons into the cell cytosol in response to lowering pH_o (fig. S10),

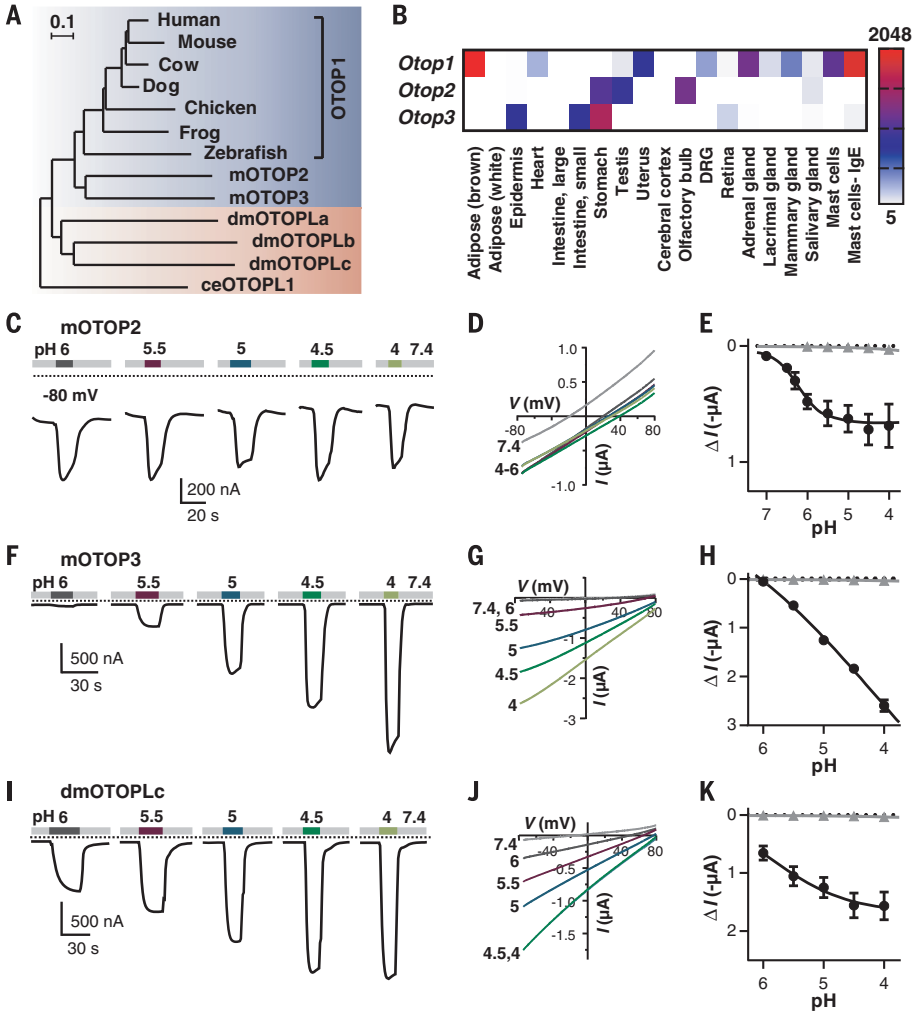


Fig. 3. An evolutionarily conserved family of genes, expressed in diverse tissues and encoding proton channels. (**A**) Maximum-likelihood phylogenetic tree from the multisequence alignment of 13 otopetrin domain proteins. Scale bar indicates amino acid substitutions per site. dm, *Drosophila melanogaster*; ce, *Caenorhabditis elegans*. (**B**) Distribution of *Otop* genes in selected murine tissues from microarray data (*16*). Scale represents expression level in arbitrary units (mean ± SEM, n = 2). (**C**, **F**, **I**) Representative traces (V_m = −80 mV) showing currents evoked in *Xenopus* oocytes expressing OTOP2, OTOP3, or dmOTOPLc in response to varying pH_o of the Na^+-free extracellular solution. (**D**, **G**, **J**) *I-V* relationship (from voltage ramps at 1 V/s) from experiments as in (C), (F), and (I). (**E**, **H**, **K**) The average current induced at V_m = −80 mV (ΔI) as a function of pH for oocytes expressing each of the channels (black circles; mean ± SEM, n = 3 to 7) and for uninjected oocytes (gray triangles, mean ± SEM, n = 3).

providing evidence that, like OTOP1, they form proton channels.

There are three genes in the genome of *Drosophila melanogaster* that encode proteins that appear to be evolutionarily related to mOTOP1 (*12*, *14*) (Fig. 3A). The transcript CG42265 encodes dmOTOPLc, a protein of 1576 amino acids that over the region of similarity bears 14.1% amino acid identity with OTOP1. Despite the modest level of conservation, when expressed in *Xenopus* oocytes, dmOTOPLc conducted large currents in response to decreasing extracellular pH, indicating that it too forms a proton channel (Fig. 3, I to K). The shallow relation between the current amplitude and pH may endow the channel with a broader dynamic range.

OTOP1 is required for the development of otoconia, calcium carbonate–based structures that sense gravity and acceleration in the vestibular system. Two mutations of *Otop1*, tilted (*tlt*) and mergulhador (*mlh*; fig. S11A), lead to vestibular dysfunction in mice (*14*). These mutations affect trafficking of the protein to the cell surface in vestibular supporting cells (*17*). Mutant channels expressed in *Xenopus* oocytes produced smaller currents but otherwise had functional properties, such as sensitivity to Zn^{2+} (fig. S11, B and C), similar to those of wild-type OTOP1. This reduction in current magnitude may contribute to the vestibular dysfunction.

Finally, we sought to determine if OTOP1 contributes to the proton current in acid-sensing taste receptor cells (*10*, *11*). We confirmed that in single-cell transcriptome data, *Otop1* was expressed in PKD2L1 cells (19 out of 19) implicated in sour transduction (*18*), whereas *Otop2* and *Otop3* were expressed in much lower amounts, and none of the three *Otop* transcripts were detected in TRPM5 cells, which lack proton currents (Fig. 4A). By immunocytochemistry, we confirmed that OTOP1 was present in taste cells in mouse circumvallate papillae that express *Pkd2l1* (Fig. 4B). To directly determine if OTOP1 contributes to the proton current in taste cells, we measured currents in taste cells from either wild-type mice or mice that were homozygous for the *tlt* mutation of *Otop1*. Mutation of *Otop1* resulted in significantly smaller proton currents than those measured in taste cells from wild-type mice (Fig. 4, C and D), over a range of H^+ concentrations (Fig. 4, E and F), indicating that OTOP1 is a component of the proton channel in taste cells. Although the contribution of proton currents to acid-sensing or sour taste behavior by mice is still speculative and complicated by contributions from multiple sensory organs and sensory receptors (*19*), the identification of OTOP1 as forming a proton channel provides a tool with which to start dissecting this system.

Our data show that the *otopetrin* genes encodes a family of ion channels that are unrelated structurally to previously identified ion channels and are highly selective for protons. Unlike Hv1, OTOP1 is only weakly sensitive to voltage. Whether, like the viral proton channel M2 (*13*), low pH gates OTOP1 is not clear. OTOP channels conduct protons at normal resting potentials and can mediate the entry of protons into cells. Most cells guard against proton entry, which is generally cytotoxic. Thus, we expect that OTOP channels are restricted to cell types that use changes in intracellular pH for cell signaling or to regulate biochemical or developmental processes. Along with a role in formation of vestibular otoconia (*14*), OTOP1 has been shown to protect mice from obesity-induced metabolic dysfunction (*15*), and it is up-regulated in dorsal root ganglion cells in response to cell damage (*20*). The knowledge that this gene family encodes proton channels can be used to understand its contribution to physiology and disease.

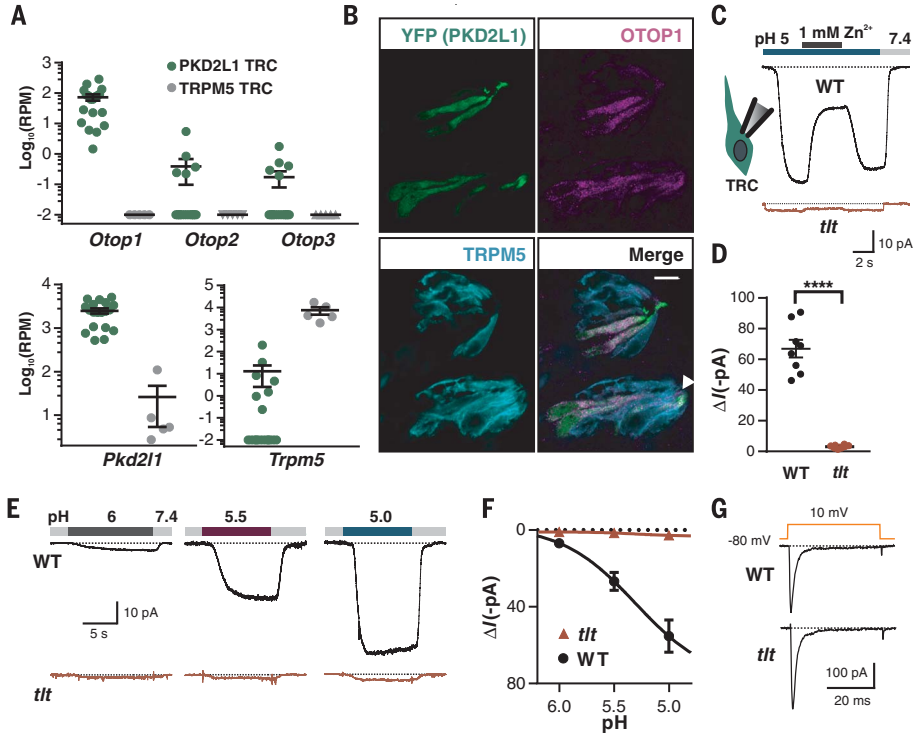


Fig. 4. Requirement of *Otop1* for the proton current in taste receptor cells. (**A**) Read counts per million (RPM) for the genes indicated from RNA-sequencing data obtained from single PKD2L1 (n = 19) or TRPM5 taste cells (n = 5). 0 RPM was adjusted to 0.01 RPM. (**B**) Confocal images showing taste buds in the circumvallate papillae from a mouse in which *Pkd2l1* drives expression of YFP, immunostained with antibodies against YFP (green), OTOP1 (magenta), and TRPM5 (cyan). Scale bar, 10 μM. Arrow indicates taste pore. (**C**) Current in response to a pH 5.0 stimulus in isolated PKD2L1 TRCs from *tlt* mutant or wild-type (WT) mice in $NMDG^+$-based solution (V_m = −80 mV). (**D**) Average data from experiments as in (C) (****P < 0.0001 by two tailed t test, n = 8 cells per genotype). (**E**) Response of PKD2L1 TRCs to $NMDG^+$-based extracellular solution of varying pH (V_m = −80 mV). (**F**) Average data from experiments as in (E). (**G**) Voltage-gated Na^+ currents in TRCs from *tlt* and wild-type mice were indistinguishable (P > 0.05, two-tailed t test).

REFERENCES AND NOTES

1. B. Hille, *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, MA, 2001).
2. W. A. Catterall, G. Wisedchaisri, N. Zheng, *Nat. Chem. Biol.* **13**, 455–463 (2017).
3. E. Gouaux, R. Mackinnon, *Science* **310**, 1461–1465 (2005).
4. T. E. Decoursey, *Physiol. Rev.* **83**, 475–579 (2003).
5. L. H. Pinto, L. J. Holsinger, R. A. Lamb, *Cell* **69**, 517–528 (1992).
6. T. E. DeCoursey, *Biophys. J.* **60**, 1243–1253 (1991).
7. I. S. Ramsey, M. M. Moran, J. A. Chong, D. E. Clapham, *Nature* **440**, 1213–1216 (2006).
8. M. Sasaki, M. Takagi, Y. Okamura, *Science* **312**, 589–592 (2006).
9. D. Morgan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18022–18027 (2009).
10. R. B. Chang, H. Waters, E. R. Liman, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22320–22325 (2010).
11. J. D. Bushman, W. Ye, E. R. Liman, *FASEB J.* **29**, 3014–3026 (2015).
12. I. Hughes *et al.*, *BMC Evol. Biol.* **8**, 41 (2008).
13. I. V. Chizhmakov *et al.*, *J. Physiol.* **494**, 329–336 (1996).
14. B. Hurle *et al.*, *Hum. Mol. Genet.* **12**, 777–789 (2003).
15. G. X. Wang *et al.*, *Diabetes* **63**, 1340–1352 (2014).
16. C. Wu *et al.*, *Genome Biol.* **10**, R130 (2009).
17. E. Kim *et al.*, *Mol. Cell. Neurosci.* **46**, 655–661 (2011).
18. A. L. Huang *et al.*, *Nature* **442**, 934–938 (2006).
19. E. R. Liman, Y. V. Zhang, C. Montell, *Neuron* **81**, 984–1000 (2014).
20. A. K. Reinhold *et al.*, *PLOS ONE* **10**, e0123342 (2015).

ACKNOWLEDGMENTS

E.R.L. and the University of Southern California have filed a provisional patent application no. 62/537,900 that claims methods of screening molecules that modulate Otopetrin-dependent ion channel activities. We thank S. Rao, L. Goggins, A. Bernanke, P. Uren, and members of the Nuzhdin lab for technical assistance; J. Bushman for assistance with electrophysiology; and D. Arnold, B. Bean, B. Herring, and R. Kramer for careful reading of the manuscript. Funding was provided by the NIH grants R01DC013741 and R21DC012747 (to E.R.L.) and R01HG006015 (to A.D.S.).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1047/suppl/DC1
Materials and Methods
Figs. S1 to S11
Table S1
References (*21*–*34*)

13 July 2017; accepted 8 January 2018
Published online 25 January 2018
10.1126/science.aao3264

MOLECULAR BIOLOGY

Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements

Bo Gu,[1] Tomek Swigut,[1] Andrew Spencley,[1,2] Matthew R. Bauer,[1] Mingyu Chung,[1] Tobias Meyer,[1] Joanna Wysocka[1,3,4,5]*

To achieve guide RNA (gRNA) multiplexing and an efficient delivery of tens of distinct gRNAs into single cells, we developed a molecular assembly strategy termed chimeric array of gRNA oligonucleotides (CARGO). We coupled CARGO with dCas9 (catalytically dead Cas9) imaging to quantitatively measure the movement of enhancers and promoters that undergo differentiation-associated activity changes in live embryonic stem cells. Whereas all examined functional elements exhibited subdiffusive behavior, their relative mobility increased concurrently with transcriptional activation. Furthermore, acute perturbation of RNA polymerase II activity can reverse these activity-linked increases in loci mobility. Through quantitative CARGO-dCas9 imaging, we provide direct measurements of cis-regulatory element dynamics in living cells and distinct cellular and activity states and uncover an intrinsic connection between cis-regulatory element mobility and transcription.

Cis-regulatory DNA elements such as promoters and long-range enhancers mediate precise spatiotemporal control of gene expression (*1*–*4*). Understanding regulatory elements' dynamics in living cells and their changes in relation to transcriptional status and cellular states is important for comprehending gene expression control. To explore these questions, we first had to address technical limitations associated with the labeling of small, nonrepetitive genomic elements. Previous strategies typically relied on inserting heterologous arrays of bacterial operator sequences that are much larger than a typical individual enhancer or promoter and, owing to their highly repetitive nature, may be subject to regulation specific for repetitive sequences (*5*–*10*). Alternatively, dCas9 (catalytically dead Cas9) imaging can facilitate RNA-guided labeling of native genomic regions (*11*), but similarly to other methods, it requires a large number of fluorescent molecules bound to the target locus at any given time to enable microscopic visualization. Consequently, in the absence of approaches allowing for highly multiplexed and uniform delivery of guide RNAs (gRNAs) to target cells, dCas9 imaging has been practically limited to the labeling of repetitive sequences (*12*–*15*).

To overcome this bottleneck, we developed a molecular assembly strategy, termed chimeric array of gRNA oligonucleotides (CARGO), that can achieve highly multiplexed gRNA delivery into single cells (diagram in Fig. 1A). With this strategy, gRNA 12- and 18-nucleotide monomers can be readily assembled in a single step with 70% and 60% efficiencies, respectively, and an even higher degree of multiplexing can be readily accomplished (fig. S1; see materials and methods and other supplementary materials for details and protocol). We then asked whether robust labeling of nonrepetitive cis-regulatory elements can be achieved in mammalian cells by combining CARGO with dCas9 imaging. To this end, we chose mouse embryonic stem cells (mESCs) and their in vitro differentiation to epiblast-like cells (mEpiLCs) as a cell fate transition model (*16*–*18*) (fig. S2A). We have previously shown that during this transition a cluster of enhancers located ~30 to 54 kb downstream from the *Fgf5* promoter is activated de novo (*17*) (fig. S2B) and that these changes are accompanied by the transcriptional induction of the *Fgf5* gene (*17*). To label this developmentally regulated enhancer cluster, we designed three CARGO arrays, each harboring 12 different gRNAs, spanning a 2-kb window immediately upstream of the first enhancer (E1) within the cluster (fig. S2B). We generated clonal mESC lines with a stably integrated, inducible *dCas9-eGFP* transgene expressed at relatively low (albeit clone-to-clone–variable) level upon induction with doxycycline (Dox) (fig. S3, A and B). As expected (*11*), dCas9-eGFP displayed previously observed nucleolar retention in the absence of gRNA and robustly labeled telomeric repeats upon transfection with telomere gRNA (Fig. 1B). Upon transfection of the three *Fgf5* enhancer CARGO arrays, one or two puncta were clearly visible in live cells to which arrays were successfully introduced, as measured by a fluorescent marker encoded by the CARGO plasmid (Fig. 1B). The labeling was specific, as confirmed by the colocalization of the dCas9-eGFP immunofluorescence and *Fgf5* DNA fluorescence in situ hybridization (FISH) signals (Fig. 1C), and efficient, as 60% to 70% of cells across different lines were consistently labeled with one or two puncta with a sufficiently high signal to background ratio (Fig. 1, D and E).

We next investigated both the precision of the dCas9 targeting and its potential interference with enhancer activation during differentiation. We performed anti-GFP (green fluorescent protein) and anti-H3K27ac (acetylation of histone H3 at lysine 27) chromatin immunoprecipitation (ChIP)–quantitative polymerase chain reaction (qPCR) analyses from dCas9-eGFP mESCs with or without a CARGO array, and from mEpiLCs derived from them through 48 hours of differentiation. Using qPCR amplicons spanning either the CARGO-array targeted region (a to c) or enhancer E1 itself (d and e) (fig. S2B), we observed that: (i) dCas9 recruitment is dependent on the array and localized to the targeted region, and (ii) enhancer activation, as indirectly measured by H3K27ac, retains its developmental dynamics and is not considerably affected by dCas9-eGFP recruitment in the vicinity (fig. S2, C and D). In agreement, *Fgf5* mRNA single-molecule FISH (smFISH) analysis of mEpiLC dCas9-eGFP cells with or without a CARGO array showed no significant differences in transcript numbers between the two populations (fig. S4A). Taken together, these results demonstrate that CARGO-dCas9 imaging provides a specific and noninvasive strategy to label functional cis-regulatory sequences in their native chromosomal context.

To follow the movement of the *Fgf5* enhancer in its active and inactive state, we performed live CARGO-dCas9 imaging in mESCs or in mEpiLCs after 48 hours of differentiation and tracked centers of labeled loci with high temporal resolution (Fig. 2A and movie S1). Visual inspection of the recorded time-lapse images revealed that a substantial fraction of *Fgf5* enhancers displayed increased mobility in mEpiLCs compared with mESCs (compare movies S2 and S3), and this was also evident when we examined individual trajectories over a fixed time interval (Fig. 2A, compare c and d). To characterize the *Fgf5* enhancer movement quantitatively, we computed the mean square displacement (MSD) for these trajectories (Fig. 2B) and extracted two parameters: the scaling exponent α and the apparent diffusion coefficient D_{app}. Time-averaged MSD (tMSD) plots show an anomalous scaling exponent of $\alpha = 0.53 \pm 0.21$ in mESCs and $\alpha = 0.51 \pm 0.26$ in mEpiLCs (Fig. 2B), suggesting subdiffusive behavior of the enhancer, similar to chromosome movement in bacteria, yeast, and B cells (*8*, *19*, *20*). Additionally, a comparable scaling exponent α (0.54 in mESCs and 0.51 in mEpiLCs) was obtained from the time- and ensemble-averaged MSD (eMSD) plots (Fig. 2B, shaded area denotes SEM), indicating that the loci movements are ergodic.

Whereas the scaling exponent extracted from our MSD measurements was similar for the two examined cell states, the apparent anomalous diffusion coefficient of the *Fgf5* enhancer exhibited a significant increase in mEpiLCs compared with ESCs (Fig. 2C and table S2). Moreover, distribution of D_{app} was unimodal in the mESC state,

[1]Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA 94305, USA. [2]Cancer Biology Program, Stanford University School of Medicine, Stanford, CA 94305, USA. [3]Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305, USA. [4]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA. [5]Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305, USA.
*Corresponding author. Email: wysocka@stanford.edu

whereas in the mEpiLC state the observed data can be better explained by the appearance of an additional "fast" population with a higher apparent anomalous diffusion coefficient D_{app} ($8.6 \times 10^{-3 \pm 0.38}$ $\mu m^2/s^{0.5}$) representing ~69% of the tracked enhancer alleles (with the remaining alleles showing "slow" behavior with D_{app} of $1.6 \times 10^{-3 \pm 0.16}$ $\mu m^2/s^{0.5}$) (Fig. 2C). These measurements are consistent with the increased mobility of the enhancer in mEpiLCs in our time-lapse images.

We next used CARGO-dCas9 to label the *Fgf5* promoter, which is transcriptionally induced during the mESC-to-mEpiLC transition [Fig. 3A, fig. S4 (for smFISH showing that labeling does not interfere with Fgf5 expression), and fig. S5, top panel (for the CARGO array position in relation to the transcription start site, TSS)]. Tracking and MSD measurements of *Fgf5* promoter movement in mESCs showed α and D_{app} values consistent with the slow subdiffusive behavior observed for the *Fgf5* enhancer in mESCs (table S2). Furthermore, an overall increase in mobility was observed for the *Fgf5* promoter in the mEpiLC state, which can also be attributed to the appearance of a fast subpopulation, with D_{app} of $26 \times 10^{-3 \pm 0.28}$ $\mu m^2/s^{0.5}$ (Fig. 3B, first row). Two other promoters that become transcriptionally induced during the mESC-to-mEpiLC transition, *Otx2* and *Oct6*, also show elevated mobility in the active state, whereas the *Dusp5* promoter, which maintains a consistently low transcriptional activity, shows unimodal distribution of

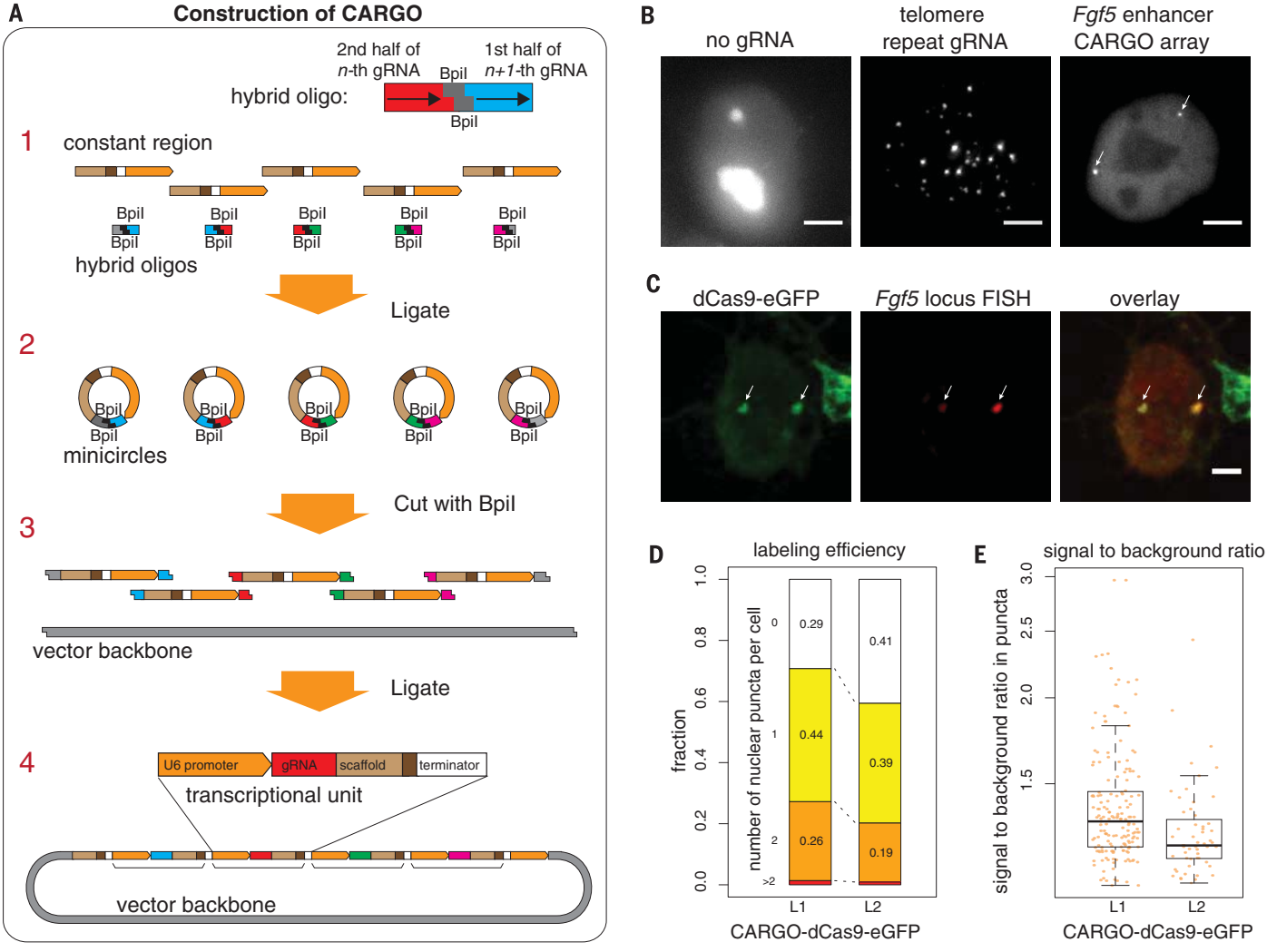


Fig. 1. CARGO-dCas9 imaging enables robust and noninvasive labeling of cis-regulatory elements in living cells. (**A**) CARGO assembly of a multiplexed gRNA array. Hybrid DNA oligonucleotides are synthesized with the first half of *n*th gRNA sequence, followed by the second half of the (*n* – 1)th gRNA separated by two BpiI restriction sites, with distinct sticky ends for each gRNA. Step 1: Synthetic DNA oligonucleotides are mixed and ligated with a permuted expression unit constant region (gRNA scaffold, Pol III termination signal, and human U6 promoter). Step 2: Resulting mini circles are cut with BpiI, exposing complementary sticky ends from different circles. Step 3: Digested products and destination vector are ligated to produce an array of gRNA expression units (shown in step 4) in a single-pot reaction. (**B**) Representative examples of dCas9 imaging of genomic loci in mESCs. (Left) No gRNA control. (Middle) Single-gRNA–targeting telomere repeats. (Right) CARGO-array–targeting *Fgf5* enhancer. Scale bars, 5 μm. (**C**) Representative images showing colocalization of the *Fgf5* enhancer CARGO dCas9-eGFP signal [as visualized by anti-eGFP (anti–enhanced GFP) immunofluorescence] with the DNA FISH signal (position of a BAC FISH probe is shown in fig. S3C). Scale bar, 2 μm. Colocalization was confirmed by Fisher's exact test: $P < 3 \times 10^{-28}$; odds ratio = 1.49×10^4 of nonrandom association between sparsely sampled dCas9 and DNA FISH image pixels. (**D** and **E**) CARGO-dCas9 locus labeling efficiency (D) and signal-to-background ratio (E) in two clonal mESC lines (L1 and L2) bearing dCas9-GFP fusion and transfected with Fgf5 enhancer CARGO arrays. In (E), the bold line at the center of each box denotes the median value; top and bottom edges of the box denote the 25th and 75th percentiles, respectively.

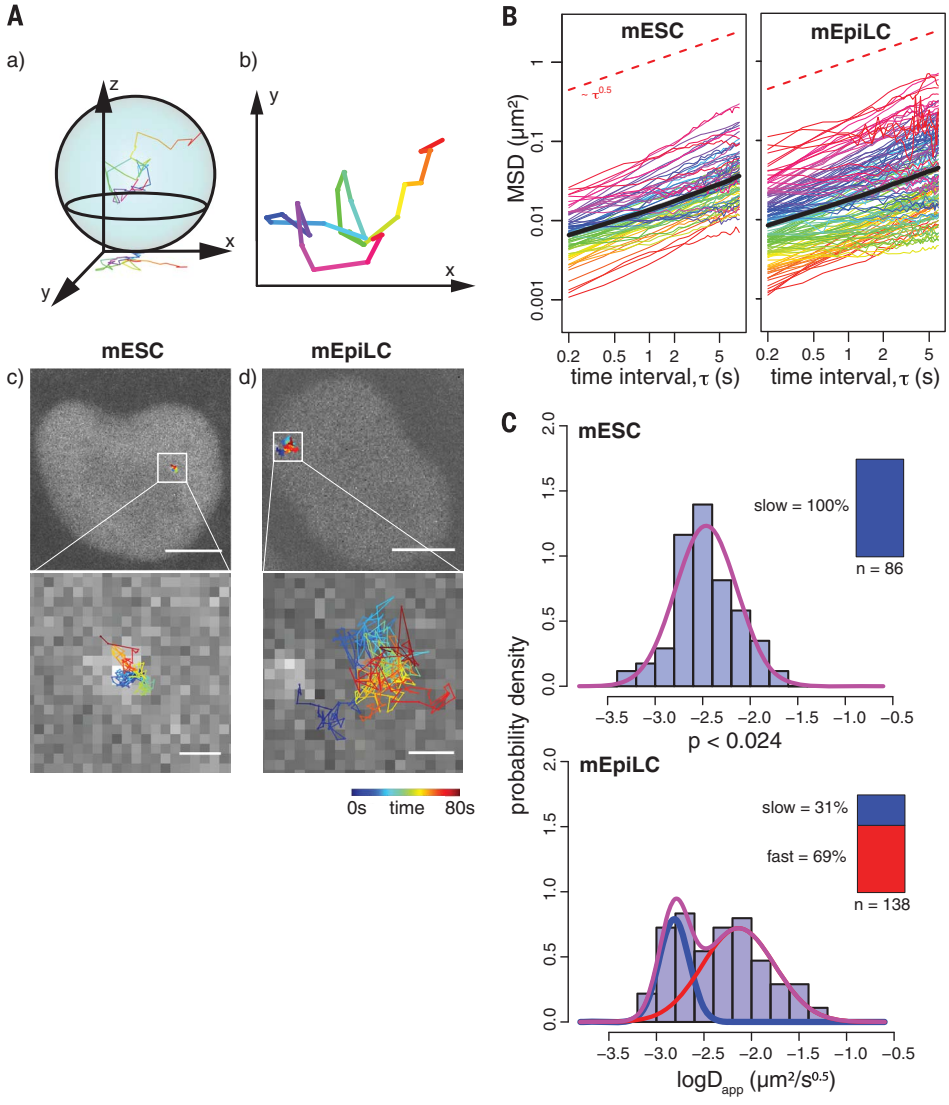


Fig. 2. Live-cell CARGO-dCas9 imaging and tracking of the *Fgf5* enhancer during differentiation of mESCs to mEpiLCs. (**A**) Live-cell two-dimensional (2D) tracking of the CARGO-dCas9–labeled *Fgf5* enhancer in mESCs and mEpiLCs. (a and b) Movies of the cell nuclei are recorded as 2D projections of the 3D movement of the labeled loci. (c and d) Representative images of a single mESC (c) or mEpiLC (d) nucleus, overlaid with recorded trajectories color-coded by time (0 to 80 s). The bottom panels show zoomed-in views of the inset areas in the top images. Scale bars, 5 μm (top); 500 nm (bottom). (**B**) Subdiffusive motion of the *Fgf5* enhancer locus. tMSD for each tracked enhancer allele (colored curves) and eMSD (bold black curve, shaded area indicates ± SEM) as a function of the time interval (τ) between observations. Ninety-one and 130 observed alleles are plotted for the mESC and mEpiLC state, respectively. The red dashed reference line has a slope of 0.5. (**C**) Appearance of a fast-moving *Fgf5* enhancer population in the mEpiLC state. Histograms of fitted apparent anomalous diffusion coefficients calculated from tMSD trajectories in mESCs or mEpiLCs, as indicated, are overlaid with fitted Gaussian mixture distribution curves in purple. Individual slow and fast components are plotted as blue and red curves, respectively. The inset bar plots indicate the number of recorded trajectories, together with the mixing proportion of slow and fast population components for each individual Gaussian. Bayesian information criterion (BIC) values for different component fittings are listed in table S1. The difference in distributions is supported by a two-sample Kolmogorov-Smirnov test.

D_{app} with slow diffusivity in both cellular states [Fig. 3, A and B; fig. S5, bottom panel; fig. S6 (for tMSD and eMSD plots); and fig. S7D]. These observations suggest that mobility of cis-regulatory elements may be linked to their activity status and that the underlying heterogeneity of this dynamics could reflect the heterogeneity of the transcriptional state. Alternatively, these mobility shifts could be explained by the global differences in chromatin properties and/or compaction associated with mESC and mEpiLC states.

To distinguish between these two possibilities, we labeled the promoter and distal superenhancer of *Tbx3*, a gene that becomes down-regulated during the mESC-to-mEpiLC transition (Fig. 3A and fig. S5, middle panel). Notably, CARGO-dCas9 imaging of the *Tbx3* promoter or superenhancer located ~90 kb away from the TSS revealed changes in mobility that were opposite to those observed at the *Fgf5* locus [Fig. 3B, fig. S6 (for tMSD and eMSD plots), and movies S4 and S5].

Analysis of the combined live-cell tracking data of all seven regulatory elements in both cellular states suggests that as a first-order approximation, the observed heterogeneity of cis-regulatory element dynamics can be better explained by a two-population model [$\Delta BIC = -143.2$; *BIC*, Bayesian information criterion] versus a single-population model, where the apparent diffusivity of the fast population is three to four times higher than that of the slow population (Fig. 3C). We compared the extracted fast versus slow population fractions for all seven tested loci in either mESCs or mEpiLCs and observed a statistically significant increase of the fast population in the transcriptionally active state (Fig. 3D and fig. S7A). However, because the fitted standard deviations of D_{app} distribution in the fast state are broad, it remains an open question whether a single D_{app} value for the fast state exists or whether there is a continuum that is dependent on the level of transcription and locus-specific effects.

The scaling exponent α was comparable (~0.5) for all examined elements in both cellular states (fig. S7B and table S2) and close to the α values obtained through measurements of chromatin movement in mammalian cells, yeast, and bacteria. These observations suggest that the subdiffusive behavior and scaling laws governing the movement of cis-regulatory elements in the interphase nuclei are similar to those of other classes of chromatin and are conserved across the tree of life. Furthermore, the comparable shape and negative dip of the time-lag–rescaled velocity autocorrelation function indicate that the viscoelastic nature of nucleoplasm and chromatin fibers can explain the subdiffusive scaling behavior of cis-regulatory elements (*8*) in both the active and inactive states (fig. S7C).

To gain more insight into the relationship between the cis-regulatory element mobility and the transcriptional status of its cognate gene at the single-cell level, we collected matched live–fixed cell data by tracking *Fgf5* enhancer movement in live mEpiLCs and measuring transcriptional status in the same cells by smFISH with the intronic and exonic *Fgf5* mRNA probes (see fig. S8 for specificity validation). We took advantage of the expression heterogeneity during differentiation and classified cells into three categories: actively transcribing (*Fgf5* intronic +, *Fgf5* exonic +),

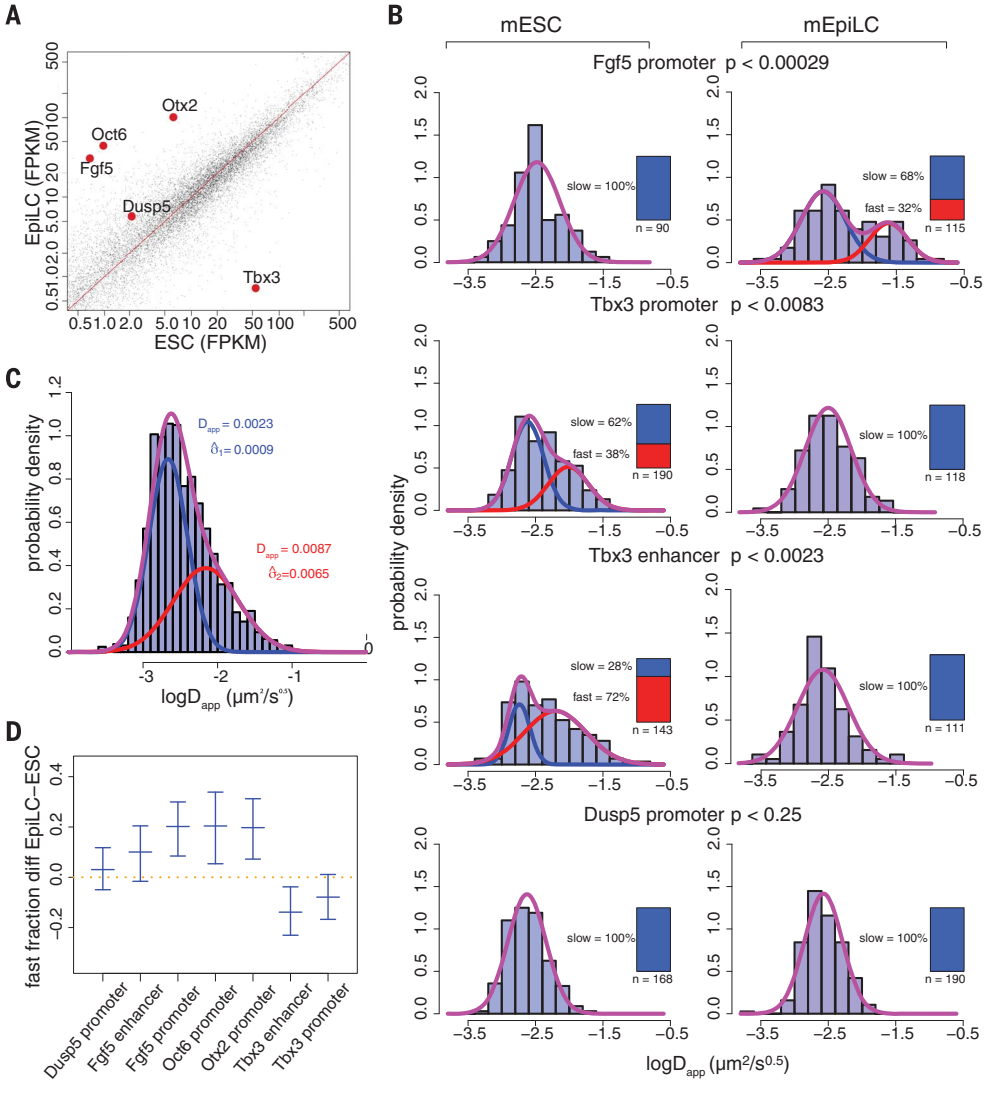


Fig. 3. Mobility of cis-regulatory elements changes with the transcriptional status of their associated genes. (**A**) Expression changes of genes whose cis-regulatory elements were analyzed by live-cell tracking. Ordinate: mean expression in mESC state; abscissa: mean expression in mEpiLC state, as measured by RNA sequencing. All genes are shown, with investigated genes highlighted in red. FPKM, fragments per kilobase of transcript per million mapped reads. (**B**) Histograms of fitted apparent anomalous diffusion coefficients from tMSD of the indicated regulatory regions in mESCs (left panels) or mEpiLCs (right panels), overlaid with fitted Gaussian mixture distributions (purple) along with slow (blue) and fast (red) components. The inset bar plots indicate the number of recorded trajectories and the individual Gaussian mixing proportion of slow and fast population components. BIC values for different component fittings are listed in table S1. Differences in the distribution between the mESC and mEpiLC states are supported by a two-sample Kolmogorov-Smirnov test. (**C**) Histogram of combined fitted apparent anomalous diffusion coefficients from tMSD of all loci in both mESC and mEpiLC states (n = 1271) overlaid with the fitted Gaussian mixture distribution (purple) along with slow (blue) and fast (red) components. Log-normal means and standard deviations of the slow and fast components (in $\mu m^2\ s^{-0.5}$) are denoted on the plot. (**D**) Difference in fractions of fast populations between the mESC and mEpiLC states. Center bars indicate the median value; upper and lower bars indicate the 95% confidence interval of the estimates.

recently active (*Fgf5* intronic −, *Fgf5* exonic +), or inactive (*Fgf5* intronic −, *Fgf5* exonic −). By analyzing the smFISH signals in relation to the *Fgf5* enhancer mobility, we observed that in cells with *Fgf5* mRNA transcripts, the enhancer explores a larger nuclear space over a given time compared with in the inactive cells, with the highest mobility observed for enhancer alleles that are actively transcribing at the time of measurement (Fig. 4, A and B). This strong association suggests a direct connection between transcriptional activity and cis-regulatory element mobility.

To directly test the relationship between the status of RNA polymerase II (Pol II) and the anomalous diffusive behavior of the *Fgf5* enhancer, we acutely (for 10 to 30 min) perturbed Pol II activity by using chemical inhibitors targeting either transcriptional elongation (DRB and flavopiridol) (*21*, *22*) or transcriptional initiation (triptolide) (*23*, *24*). Tracking *Fgf5* enhancer mobility revealed that inhibition of either initiation or elongation results in a significant reduction of D_{app}, both in bulk measurements of eMSD across all cells (Fig. 4C) and at the single-cell level, as measured by tMSD in matched single cells before and after the treatment (Fig. 4D). Notably, alleles showing higher D_{app} values in untreated cells (and thus representing the fast subpopulation) were the most affected by the Pol II inhibition. Similar decreases in D_{app} upon Pol II inhibition were also observed for the *Tbx3* promoter in mESCs (fig. S9). Together, our observations indicate that the increased nuclear mobility of cis-regulatory elements is directly coupled to the Pol II activity and transcriptional status of their cognate genes.

The CARGO assembly method described here has a wide application potential beyond imaging, in techniques such as CRISPR interference, CRISPR activation, and in simultaneous editing of multiple genomic regions, where multiplexed recruitment of dCas9 or Cas9 to genomic loci of interest is required or desired. Through CARGO-dCas9 visualization and quantitative measurement of cis-regulatory element dynamics in distinct activity states, we uncovered an unexpected relationship between the mobility of cis-regulatory elements and local transcriptional activity. What may be the underlying causes of the observed transcription-coupled increases in mobility? Although the exact formulations vary with different polymer models (*25*), D_{app} generally rises with increased molecular agitation (e.g., absolute temperature in the simplest form but also nonthermal sources of agitation) and with decreased friction and/or viscosity and polymer size. Several lines of evidence suggest that the increase in D_{app} may be due to a nonthermal molecular motion generated by an energy-dissipating biochemical process (*26*). The main candidate for such a process in our studies is the activity of RNA Pol II—a major molecular motor that translocates through the chromatin and dissipates energy via pyrophosphate release—although transcription-coupled adenosine triphosphate (ATP)–dependent chromatin remodelers may also contribute to the local movement. Collectively, we propose that

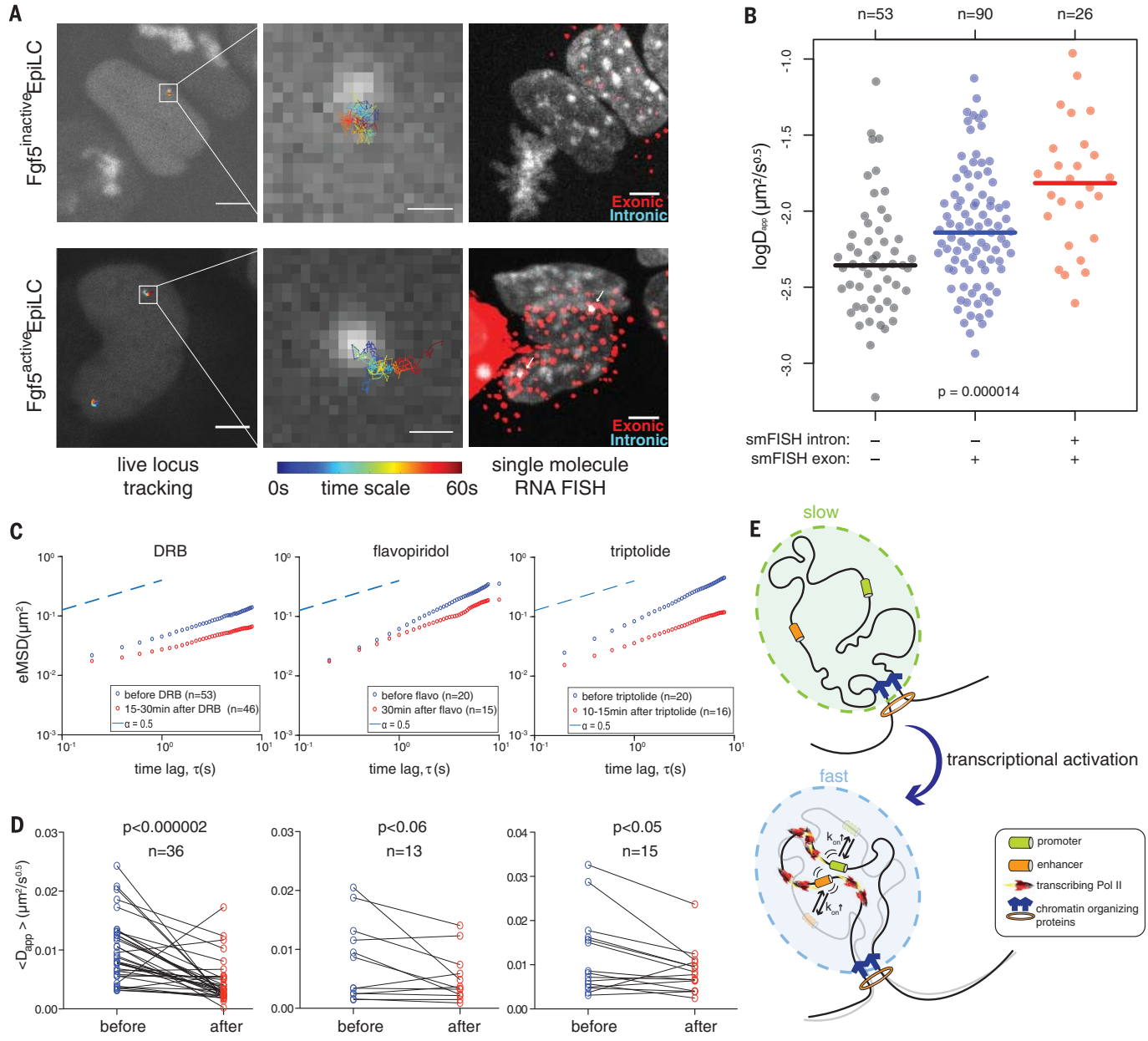


Fig. 4. Inhibition of RNA polymerase II reverses activity-associated changes in enhancer mobility. (**A**) Live-cell locus tracking correlation with multiplexed single-molecule RNA FISH. Representative images of a cell with an inactive (top) or active (bottom) *Fgf5* locus are shown. (Left) Snapshot of live-cell imaging in grayscale overlaid with fitted *Fgf5* enhancer trajectories color-coded by frame number from 1 to 300 (corresponding to time 200 ms to 60 s). Scale bars, 5 μm. (Middle) Magnified view of the highlighted regions in the left panel. Scale bars, 500 nm. (Right) Multiplexed smFISH of *Fgf5* mRNA from the same cells. Gray channel: DAPI (4′,6-diamidino-2-phenylindole) staining for cell nucleus; red channel: smFISH probe targeting the first exon of *Fgf5* mRNA (the mCherry marker in the CARGO array is also visible); blue channel: smFISH probe targeting the first intron of *Fgf5* mRNA. Colocalized intronic and exonic smFISH signals are highlighted by arrows. Scale bars, 5 μm. (**B**) Mobility of the *Fgf5* enhancer correlates with nascent transcription of the *Fgf5* locus at the single-cell level. Cells were binned into three groups according to the transcription status of the *Fgf5* locus, as measured by multiplexed smFISH and indicated at the bottom. Individual dots represent apparent anomalous diffusion coefficients extracted from the corresponding live-imaging tMSD data. Statistical significance is supported by a Kruskal-Wallis test. (**C**) Increased mobility of the *Fgf5* enhancer in mEpiLCs is reversed by Pol II inhibition. eMSD of the *Fgf5* enhancer trajectories in mEpiLCs is shown before (blue circles) and after (red circles) treatment with DRB, flavopiridol, and triptolide, as indicated. (**D**) Increased mobility of the *Fgf5* enhancer in mEpiLCs is reversed by Pol II inhibition at the single-cell level. Anomalous diffusion coefficient of the *Fgf5* enhancer is shown for the same cells before and after the corresponding drug treatment. Differences are supported by a paired Wilcoxon test, as indicated by *P* values in the plots. (**E**) The stirring model provides an explanation for observed transcription-coupled changes in the mobility of cis-regulatory elements. The ground state (slow) is characterized by subdiffusive behavior with low apparent diffusivity governed by thermal forces. The activated (fast) state is characterized by an increased apparent diffusivity, which may be due to nonthermal agitation by transcribing RNA Pol II and/or its associated ATPases. Under the assumption that the radius of a local 3D chromosomal domain remains relatively invariant in the slow and fast states, elevated mobility of cis-regulatory elements would lead to decreased time to the first encounter between distally located enhancer and promoter regions, resulting in an increased enhancer-promoter contact frequency.

ongoing transcription may provide a source of nonthermal molecular agitation that can "stir" the chromatin within the local chromosomal domain, leading to an increase in anomalous D_{app} (Fig. 4E). We will hereafter refer to this hypothesis as the "stirring model."

The stirring model may have implications for transcription regulation: Under the assumption that the radius of the local chromosomal domain [such as a topologically associated domain (TAD)] does not substantially change in the examined cell state(s), the time to the first encounter between distally located enhancer and promoter regions should decrease along with the increased mobility within the domain. In other words, enhancer-promoter contact frequencies may increase upon transcriptional activation due to the increased probability of the stochastic encounters within the TAD, rather than due to the formation of stable enhancer-promoter loops (Fig. 4E). This type of mechanism could provide a positive-feedback loop facilitating gene expression robustness once transcription is initiated, as increased mobility would boost enhancer-promoter contact frequencies, in turn leading to more transcription.

REFERENCES AND NOTES

1. H. K. Long, S. L. Prescott, J. Wysocka, *Cell* **167**, 1170–1187 (2016).
2. M. Levine, *Curr. Biol.* **20**, R754–R763 (2010).
3. J. Dekker, L. Mirny, *Cell* **164**, 1110–1121 (2016).
4. S. Remeseiro, A. Hörnblad, F. Spitz, *Dev. Biol.* **5**, 169–185 (2016).
5. H. J. Nielsen, Y. Li, B. Youngren, F. G. Hansen, S. Austin, *Mol. Microbiol.* **61**, 383–393 (2006).
6. C. C. Robinett *et al*., *J. Cell Biol.* **135**, 1685–1700 (1996).
7. W. F. Marshall *et al*., *Curr. Biol.* **7**, 930–939 (1997).
8. J. S. Lucas, Y. Zhang, O. K. Dudko, C. Murre, *Cell* **158**, 339–352 (2014).
9. J. Vazquez, A. S. Belmont, J. W. Sedat, *Curr. Biol.* **11**, 1227–1239 (2001).
10. S. H. Mitsuda, N. Shimizu, *PLOS ONE* **11**, e0161288 (2016).
11. B. Chen *et al*., *Cell* **155**, 1479–1491 (2013).
12. B. Chen *et al*., *Nucleic Acids Res.* **44**, e75 (2016).
13. H. Ma *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3002–3007 (2015).
14. H. Ma *et al*., *Nat. Biotechnol.* **34**, 528–530 (2016).
15. Y. Fu *et al*., *Nat. Commun.* **7**, 11707 (2016).
16. K. Hayashi, H. Ohta, K. Kurimoto, S. Aramaki, M. Saitou, *Cell* **146**, 519–532 (2011).
17. C. Buecker *et al*., *Cell Stem Cell* **14**, 838–853 (2014).
18. T. Kalkan, A. Smith, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130540 (2014).
19. S. C. Weber, A. J. Spakowitz, J. A. Theriot, *Phys. Rev. Lett.* **104**, 238102 (2010).
20. G. G. Cabal *et al*., *Nature* **441**, 770–773 (2006).
21. S.-H. Chao *et al*., *J. Biol. Chem.* **275**, 28345–28348 (2000).
22. K. Yankulov, K. Yamashita, R. Roy, J.-M. Egly, D. L. Bentley, *J. Biol. Chem.* **270**, 23922–23925 (1995).
23. D. V. Titov *et al*., *Nat. Chem. Biol.* **7**, 182–188 (2011).
24. S. Vispé *et al*., *Mol. Cancer Ther.* **8**, 2780–2790 (2009).
25. Y. Zhang, O. K. Dudko, *Annu. Rev. Biophys.* **45**, 117–134 (2016).
26. S. C. Weber, A. J. Spakowitz, J. A. Theriot, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7338–7343 (2012).

ACKNOWLEDGMENTS

We thank J. Theriot, J. Ferrell, and E. Calo for comments on the manuscript; Y. Fan for assistance with single-particle tracking (IDL tracking package) matlab script modification; R. Greenberg for coining the CARGO acronym; R. Srinivasan for contributing to the initial phases of CARGO assembly optimization; and members of the Wysocka and Meyer laboratories for discussions. **Funding:** This work was supported in part by the Howard Hughes Medical Institute, NIH R01 grant GM112720-01 and Ludwig Institute Funds (to J.W.), R35GM127026 and S10OD018073 (to T.M.), and a Henry Fan Stanford Graduate Fellowship (to B.G.). **Author contributions:** B.G., T.S., and J.W. conceived and designed the study; B.G. performed experiments with help from M.R.B.; A.S. conducted and analyzed the ChIP-qPCR assay; M.C. assisted with smFISH matlab script modification and implementation; T.M. and T.S. provided critical advice on experimental designs, data analyses, and data interpretations; J.W. supervised the project; and B.G. and J.W. wrote the manuscript with input from all coauthors. **Competing interests:** B.G., T.S., and J.W. have filed a U.S. provisional patent application relating to CARGO methodology. All other authors declare no competing interests. **Data and materials availability:** All live-cell time-lapse images will be provided upon request.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1050/suppl/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S3
References (*27*–*35*)
Movies S1 to S5
Additional Supplemental Script
Additional Protocol

7 July 2017; accepted 16 January 2018
Published online 25 January 2018
10.1126/science.aao3136

NEUROIMMUNOLOGY

β_2-adrenergic receptor–mediated negative regulation of group 2 innate lymphoid cell responses

Saya Moriyama,[1] Jonathan R. Brestoff,[1,2] Anne-Laure Flamar,[1] Jesper B. Moeller,[1,3] Christoph S. N. Klose,[1] Lucille C. Rankin,[1] Naomi A. Yudanin,[1] Laurel A. Monticelli,[1] Gregory Garbès Putzel,[1] Hans-Reimer Rodewald,[4] David Artis[1]*

The type 2 inflammatory response is induced by various environmental and infectious stimuli. Although recent studies identified group 2 innate lymphoid cells (ILC2s) as potent sources of type 2 cytokines, the molecular pathways controlling ILC2 responses are incompletely defined. Here we demonstrate that murine ILC2s express the β_2-adrenergic receptor (β_2AR) and colocalize with adrenergic neurons in the intestine. β_2AR deficiency resulted in exaggerated ILC2 responses and type 2 inflammation in intestinal and lung tissues. Conversely, β_2AR agonist treatment was associated with impaired ILC2 responses and reduced inflammation in vivo. Mechanistically, we demonstrate that the β_2AR pathway is a cell-intrinsic negative regulator of ILC2 responses through inhibition of cell proliferation and effector function. Collectively, these data provide the first evidence of a neuronal-derived regulatory circuit that limits ILC2-dependent type 2 inflammation.

The type 2 inflammatory response is a highly conserved module of the innate and adaptive immune system that is elicited after exposure to infectious and environmental triggers, such as helminth infections, allergens, venoms, and other stimuli (*1*–*3*). The hallmarks of type 2 responses include activation of T helper 2 (T_H2) cells and release of type 2 cytokines, such as interleukin-4 (IL-4), IL-5, IL-9, and IL-13; production of immunoglobulin E (IgE); and the activation of multiple effector cells, such as basophils, mast cells, and eosinophils. This cascade of immunologic events is associated with increased mucus production from goblet cells and induction of smooth-muscle contractility that contributes to the elimination of pathogens or allergic stimuli (*4*, *5*). Recent studies identified group 2 innate lymphoid cells (ILC2s) as a potent source of type 2 cytokines that contribute to the type 2 inflammatory responses (*5*–*11*). Although considerable advances have been made to define the cytokines, growth factors, and environmental stimuli that trigger ILC2 responses, the regulatory mechanisms that constrain ILC2 responses and type 2 inflammation at mucosal sites remain incompletely understood.

Mucosal and lymphoid tissues are highly innervated (*12*–*14*), and immune cells—including mast cells, macrophages, ILCs, and T cells—can be regulated by neuronal-derived bioactive molecules in certain circumstances (*15*–*23*). However, our knowledge of the complex interactions between the nervous system and ILCs in the context of type 2 inflammation is still limited. RNA sequencing (RNA-seq) of sorted murine ILC subsets revealed that small intestinal (SI) ILC2s exhibited higher levels of the β_2AR gene (*Adrb2*) mRNA expression compared to SI ILC3s, but lower levels of other adrenergic receptors (Fig. 1A). Consistent with these results, ILC2s sorted from SI lamina propria (SILP), gut-draining mesenteric lymph node (mLN), colon LP, and lung exhibited higher levels of *Adrb2* mRNA expression compared to SILP ILC3s (Fig. 1B and fig. S1A). ILC2s from mesenteric white adipose tissue exhibited lower levels of *Adrb2* mRNA expression compared to ILC2s from other tissue. We also detected *ADRB2* mRNA in ILC2s and $CD4^+$ T cells sorted from human lung and peripheral blood mononuclear cells (Fig. 1C). These results show that the β_2AR gene is expressed in murine and human ILC2s.

Examination of the localization of ILC2s (defined as $KLRG1^+Bcl11b^+CD3\varepsilon^-NKp46^-$) and adrenergic neurons [marked by tyrosine hydroxylase (TH) expression] in the SI microenvironment revealed that ILC2s were found in close proximity to TH^+ neurons in the villi and submucosa but not in the muscularis (Fig. 1, D and E, and fig. S1B). To further examine this, we generated *Il13*–fate mapping mice by crossing *Il13*^{cre} and Ai14 (tdTomato) mice and found that ILC2s (defined as $KLRG1^+tdTomato^+CD3\varepsilon^-$) colocalize with TH^+ neurons (Fig. 1F and fig. S1, C and D). Notably, the highest expression levels of genes encoding enzymes involved in norepinephrine synthesis, *Th* and *Dbh* (dopamine β-hydroxylase), were found in the parenchyma fraction in the SI (Fig. 1, G and H), which collectively suggests that ILC2s may respond to neuronal-derived norepinephrine, a β_2AR ligand. ILC2s were also observed in the interfollicular region (*24*), paracortical region, and the medulla of the draining mLN, which are anatomical sites where sympathetic adrenergic neurons are abundant (*12*, *13*) (Fig. 1, I to K, and fig. S1, E and F).

To analyze the effect of β_2AR signaling on ILC2 development at steady state, β_2AR-sufficient ($Adrb2^{+/+}$) and β_2AR-deficient ($Adrb2^{-/-}$) mice were analyzed. $Adrb2^{+/+}$ and $Adrb2^{-/-}$ mice had comparable numbers of ILC2 progenitors (ILC2Ps) (*25*) in the bone marrow (BM, fig. S2A). Additionally, $Adrb2^{+/+}$ and $Adrb2^{-/-}$ mice had similar frequencies of all ILC subsets in mLNs and in SILP (fig. S2, B to D), indicating that β_2AR deficiency does not result in impaired ILC2 development or homeostasis at steady state.

Next, we examined whether β_2AR signaling regulates ILC2 responses and type 2 inflammation after exposure to the gastrointestinal helminth *Nippostrongylus brasiliensis*. This infection induces potent ILC2 responses that play an important role in the expulsion of the parasite through production of IL-5 and IL-13 (*10*, *26*, *27*). After infection, $Adrb2^{-/-}$ mice exhibited increased frequencies and numbers of ILC2s and increased numbers of IL-13–producing ILC2s, but comparable frequencies of ILC1s and ILC3s, in mLNs and in SILP compared to $Adrb2^{+/+}$ mice (Fig. 2, A and B, and fig. S2E). The expression of inflammatory cytokines encoded by genes *Il25*, *Il33*, and *Tslp*, which stimulate ILC2s, was comparable between $Adrb2^{+/+}$ and $Adrb2^{-/-}$ mice in the SI after infection (fig. S2F). However, exaggerated eosinophilia and goblet cell hyperplasia (Fig. 2, C and D) and reduced worm burdens (Fig. 2E) were observed in $Adrb2^{-/-}$ mice. Of note, the amount of norepinephrine in the SI was comparable at 0, 4, and 7 days postinfection (fig. S2G), suggesting that expression of this ligand is not altered during infection. Notably, although ILC2s still expressed *Adrb2* after infection, expression was at a lower level than observed in naive mice (fig. S2H), suggesting that ILC2-intrinsic β_2AR expression is dynamically regulated in response to inflammatory cues within the tissue microenvironment. Taken together, these results indicate that the β_2AR pathway may be an important negative regulator of ILC2 responses.

Given that β_2AR-deficient mice had exaggerated ILC2 responses, we next sought to test whether treatment with a β_2AR agonist, clenbuterol, would inhibit ILC2 responses and dampen type 2 inflammation. Although the proportion of ILC subsets was not changed by β_2AR agonist treatment at steady state (fig. S2I), β_2AR agonist–treated wild-type C57BL/6 (B6) mice exhibited significantly fewer ILC2s in the mLNs compared to vehicle-treated mice after *N. brasiliensis* infection (Fig. 2F). Further, agonist treatment was associated with significantly fewer IL-5– and IL-13–producing ILC2s, reduced eosinophilia, and diminished goblet cell responses (Fig. 2, G to I, and fig. S2J). Consistent with these results, agonist-treated mice had significantly higher worm burdens (Fig. 2J). We also treated the mice with another β_2AR agonist, salmeterol, during *N. brasiliensis* infection and found reduced

[1]Jill Roberts Institute for Research in Inflammatory Bowel Disease, Joan and Sanford I. Weill Department of Medicine, Department of Microbiology and Immunology, Weill Cornell Medicine, Cornell University, New York, NY 10021, USA. [2]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA. [3]Department of Molecular Medicine, University of Southern Denmark, Odense, Denmark. [4]Division of Cellular Immunology, German Cancer Research Center (DKFZ), Heidelberg, Germany.
***Corresponding author. Email: dartis@med.cornell.edu**

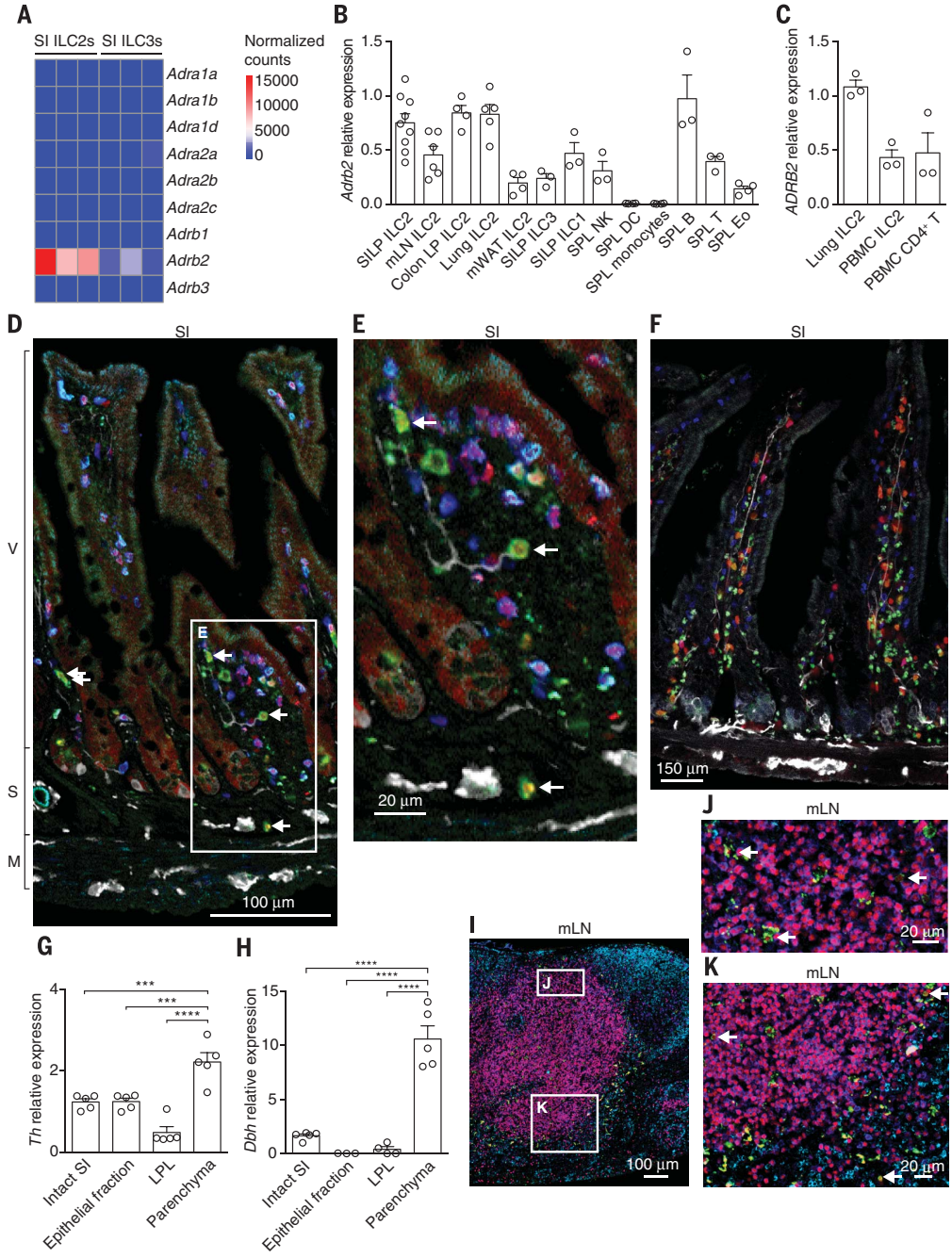


Fig. 1. *Adrb2* expression in ILCs and localization of ILC2s in gut-associated tissues. (**A**) Normalized counts of adrenergic receptor expression in SI ILC2s and ILC3s by RNA-seq are shown as a heat map. *Adra* genes encode α-ARs. (**B**) Relative expression of *Adrb2* in immune cells sorted from indicated tissues. SILP, mLN, and colon LP ILC2s were sorted as KLRG1$^+$CD127$^+$CD90$^+$Lineage negative (Lin$^-$)CD45$^+$, and lung and mesenteric white adipose tissue (mWAT) ILC2s were sorted as ST2$^+$CD127$^+$CD90$^+$Lin$^-$CD45$^+$. SILP ILC3s were sorted as CCR6$^+$CD127$^+$CD90$^+$Lin$^-$CD45$^+$. SPL, spleen; NK, natural killer cells; DC, dendritic cells; Eo, eosinophils. Data represent mean ± SEM of five pooled experiments. (**C**) Relative expression of *ADRB2* in CRTH2$^+$CD127$^+$Lin$^-$CD45$^+$ ILC2s and CD4$^+$ T cells sorted from human lung and peripheral blood mononuclear cells (PBMCs). Data represent mean ± SEM of two pooled experiments. Each circle represents data from one donor. (**D** and **E**) Representative sections of SI from a *Bcl11b*-tdTomato (red) reporter mouse stained for KLRG1 (green), CD3ε (blue), NKp46 (cyan), and TH (white). V, villi; S, submucosa; M, muscularis. Arrows show ILC2s. Data are representative of two experiments. (**F**) Representative section of SI from a *Il13*–fate mapping (red) mouse stained for KLRG1 (green), CD3ε (blue), and TH (white). Data are representative of two experiments. (**G** and **H**) Relative expression of *Th* and *Dbh* in intact SI tissue and in the indicated fractions of SI. LPL, lamina propria leukocytes. ***$P < 0.001$ and ****$P < 0.0001$ by one-way analysis of variance (ANOVA) with Dunnett's multiple comparison. (**I** to **K**) Representative section of mLN from a *Bcl11b*-tdTomato (red) reporter mouse stained for KLRG1 (green), CD3ε (blue), and IgD (cyan). Arrows show ILC2s. Data are representative of two experiments.

anti-helminth responses (fig. S2K), indicating that both β_2AR agonists have an inhibitory effect on type 2 inflammation.

Because type 2 cytokine production from ILC2s was reduced after simultaneous exposure to a β_2AR agonist and helminth infection in vivo (Fig. 2G and fig. S2J), we next sought to test whether ILC2 effector function is regulated by short-term β_2AR stimulation in vitro. SILP cells from *Il13*–YFP (yellow fluorescent protein) reporter mice were cultured with IL-33 for 4 hours in the presence of a β_2AR agonist or control vehicle. Agonist-treated cells exhibited lower frequencies of YFP-expressing ILC2s compared to vehicle-treated controls (fig. S2L), suggesting that signaling through β_2AR negatively regulates acute activation and IL-13 production from ILC2s.

In addition to ILC2s, β_2AR agonists can alter the T_H1/T_H2 cell balance by inhibiting the development of T_H1 cells (*28*). Therefore, we examined whether β_2AR stimulation by the agonist could inhibit ILC2 responses and inflammation in the absence of CD4$^+$ T cells. After treatment of B6 mice with a CD4–depleting monoclonal antibody (mAb) and the β_2AR agonist, fewer ILC2s were observed after *N. brasiliensis* infection than were observed in vehicle-treated mice (Fig. 2K). Additionally, mice treated with both anti-CD4 mAb and agonist exhibited fewer cytokine-producing ILC2s and increased worm burdens compared to vehicle-treated mice (Fig. 2, L and M). Together, these data indicate that β_2AR agonist–mediated inhibition of ILC2 responses and inflammation occurs in the absence of CD4$^+$ T cells.

To further analyze the effects of β_2AR deletion on ILCs, we crossed *Adrb2-flox* (*Adrb2$^{f/f}$*) mice with transgenic mice expressing a Cre recombinase in the *Il7r* locus (*Il7r$^{cre/+}$*) to generate β_2AR conditional knockout mice. After *N. brasiliensis* infection together with anti-CD4 mAb treatment, *Il7r$^{cre/+}$ Adrb2$^{f/f}$* mice exhibited increased ILC2 frequencies and reduced

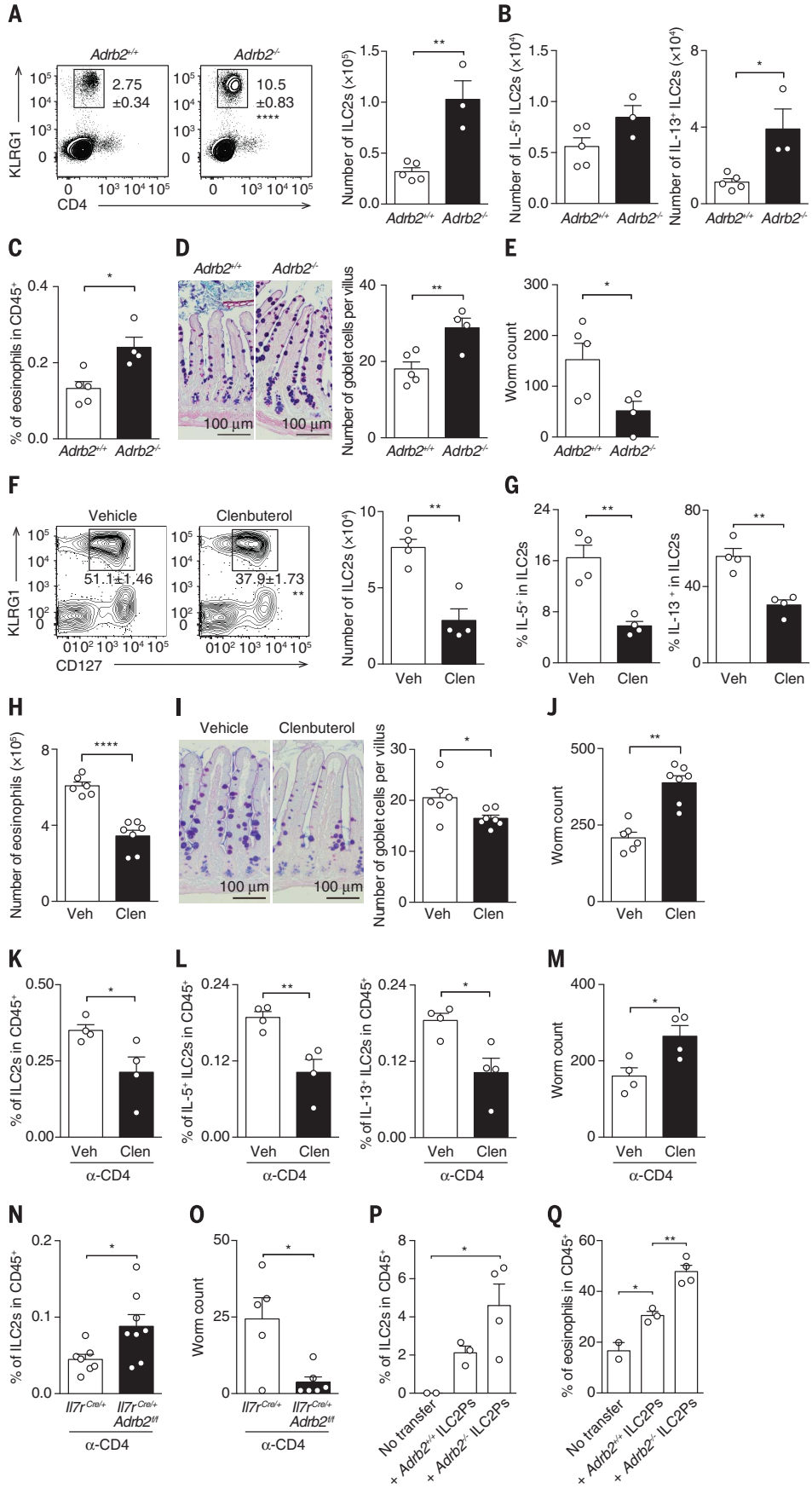


Fig. 2. Regulation of anti-helminth responses by β_2AR-deficiency or agonist treatment. (**A** to **E**) *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ mice were analyzed 4 days [(A) and (B)] and 7 days [(C) to (E)] after *N. brasiliensis* infection. Shown are (A) flow cytometry plots of Lin$^-$CD45$^+$ cells and enumeration of ILC2 percentages and numbers; (B) cytokine production; and (C) Siglec F$^+$SSC^{hi} eosinophil percentages in mLNs. Also shown are (D) representative SI sections with periodic acid–Schiff (PAS)–Alcian blue staining and enumeration of goblet cell numbers and (E) worm burdens in SI. Data are representative of three experiments. (**F** to **J**) B6 mice treated with β_2AR agonist clenbuterol (Clen) or vehicle (Veh) were analyzed 7 days after *N. brasiliensis* infection. Shown are (F) flow cytometry plots of Lin$^-$CD45$^+$ cells and enumerations of ILC2 percentages and numbers; (G) cytokine production; and (H) Siglec F$^+$SSC^{hi} eosinophil numbers in mLNs. Also shown are (I) representative SI sections with PAS–Alcian blue staining and enumeration of goblet cell numbers and (J) worm burdens in SI. Data are representative of three experiments. (**K** to **M**) B6 mice treated with anti-CD4 (α-CD4) mAb together with clenbuterol or vehicle were analyzed 10 days after *N. brasiliensis* infection. Shown are (K) enumerations of ILC2 percentages, (L) cytokine production, and (M) worm burdens in SI. Data are representative of two experiments. (**N** and **O**) *Il7r*$^{cre/+}$ and *Il7r*$^{cre/+}$ *Adrb2*$^{f/f}$ mice treated with α-CD4 mAb were analyzed 10 days after *N. brasiliensis* infection. Shown are (N) enumerations of ILC2 percentages and (O) worm burdens in SI. Data are pooled from two experiments. For panels (A) to (O), each circle represents data from one mouse, the numbers in flow cytometry plots represent mean ± SEM in each gate, and bar graphs represent mean ± SEM. *P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001 by unpaired two-tailed Student's t test. (**P** and **Q**) *Rag2*$^{-/-}$ *Il2rg*$^{-/-}$ mice reconstituted with ILC2Ps from *Adrb2*$^{+/+}$ or *Adrb2*$^{-/-}$ mice were analyzed 7 days after *N. brasiliensis* infection. Shown are enumerations of (P) ILC2 percentages and (Q) Siglec F$^+$SSC^{hi} eosinophil percentages in SI. Each circle represents data from one mouse. Bar graphs represent mean ± SEM. Data are representative of three experiments. *P < 0.05 and **P < 0.01 by one-way ANOVA with Dunnett's multiple comparison.

worm burdens compared to control *Il7r*$^{cre/+}$ mice (Fig. 2, N and O). β_2AR agonist treatment of *Il7r*$^{cre/+}$ *Adrb2*$^{f/f}$ mice did not induce significant changes in eosinophilia and worm burden after helminth infection (fig. S2, M and N). Furthermore, irradiated *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ mice reconstituted with *Adrb2*$^{+/+}$ BM cells had similar worm burdens after helminth infection (fig. S2O). These results indicate that β_2AR signaling on CD127 (IL-7 receptor α, IL-7Rα)–expressing CD4$^-$ hematopoietic cells, including ILC2s, is important during anti-helminth responses but not essential for radio-resistant cells and CD127-negative cells.

To further address the importance of β_2AR signaling directly on ILC2s, we reconstituted ILC-deficient *Rag2*$^{-/-}$ *Il2rg*$^{-/-}$ mice with *Adrb2*$^{+/+}$ or *Adrb2*$^{-/-}$ ILC2s by transferring ILC2Ps from *Adrb2*$^{+/+}$ or *Adrb2*$^{-/-}$ mice and then infecting the mice with *N. brasiliensis*. After 7 days of infection, there was a trend toward increased SI ILC2s, and we observed increased eosinophilia in *Adrb2*$^{-/-}$ ILC2–reconstituted mice compared to *Adrb2*$^{+/+}$ IL7C2–reconstituted mice (Fig. 2, P and Q), further supporting the importance of β_2AR on ILC2s in controlling anti-helminth responses.

Because lung ILC2s exhibited high levels of *Adrb2* mRNA expression (Fig. 1B) and lung tissue is also highly innervated (*15*, *19*), we examined if β_2AR signaling controls ILC2 responses in the lung as it does in the intestine. After inoculation, *N. brasiliensis* larvae migrate to and induce inflammation in the lung before reaching the intestine. *Adrb2*$^{-/-}$ mice exhibited increased frequencies of ILC2s in the lung compared to *Adrb2*$^{+/+}$ mice after infection (Fig. 3A). As an additional approach to induce lung inflammation, we administered the alarmin IL-33 intranasally and observed increased frequencies of ILC2s in *Adrb2*$^{-/-}$ mice compared to that in *Adrb2*$^{+/+}$ mice (Fig. 3B). Further, ILC2 frequencies and cytokine production after IL-33 administration were inhibited by β_2AR agonist treatment (Fig. 3, C and D). Similarly, when the mice received extract of the fungal allergen *Alternaria alternata* intranasally, β_2AR agonist–treated mice had reduced frequencies of ILC2s in the lungs than did vehicle-treated mice (Fig. 3E). In addition, *Il7r*$^{cre/+}$ *Adrb2*$^{f/f}$ mice exhibited increased ILC2 frequencies and cytokine production compared to *Il7r*$^{cre/+}$ mice after the *Alternaria* extract administration together with anti-CD4 mAb treatment (Fig. 3, F and G). Collectively, these results indicate that β_2AR signaling is an evolutionarily conserved regulatory pathway serving to dampen ILC2 responses against diverse inflammatory stimuli at multiple mucosal barrier surfaces.

These data provoke the hypothesis that the signaling through β_2AR on ILC2s negatively regulates ILC2 responses and type 2 inflammation after exposure to helminth infection and allergens. To investigate the mechanisms underlying this effect, we performed RNA-seq analysis on ILC2s sorted from *N. brasiliensis*–infected *Adrb2*$^{+/+}$ mice with or without β_2AR agonist treatment. Although ILC2s from both groups exhibited similar expression levels of genes encoding ILC2-associated transcription factors (*Id2*, *Gata3*, and *Rora*) and cytokine receptors (*Il1rl1*, *Crlf2*, and *Il4ra*) (Fig. 4A), we also observed significant down-regulation of a number of genes in the agonist-treated groups compared to that in controls. Gene set enrichment analysis (GSEA) of the down-regulated genes revealed significant enrichment in gene ontology (GO) terms associated with the cell cycle and cell proliferation (Fig. 4, B and C, and fig. S3A). Taken together with the reduced ILC2 numbers observed in these mice after infection (Fig. 2F), these results collectively suggest a role for β_2AR signaling in limiting the proliferation and accumulation of ILC2s.

To directly test whether β_2AR signaling regulates ILC2 proliferation in vivo, we analyzed the expression of the proliferation marker Ki67 during *N. brasiliensis* infection in mice treated with the β_2AR agonist. After treatment, ILC2s from these mice exhibited reduced frequencies of Ki67-expressing ILC2s, but comparable frequencies of apoptotic ILC2s, compared to those of vehicle-treated control mice after infection (Fig. 4D and fig. S3B), suggesting that β_2AR stimulation suppresses ILC2 proliferation but does not regulate apoptosis. Further, this regulation was not restricted to *N. brasiliensis*, because it was also observed after infection with another gastrointestinal helminth, *Heligmosomoides polygyrus bakeri* (Fig. 4E). In addition, when *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ mice were injected with IL-33 intraperitoneally, the frequencies of Ki67$^+$ ILC2s were

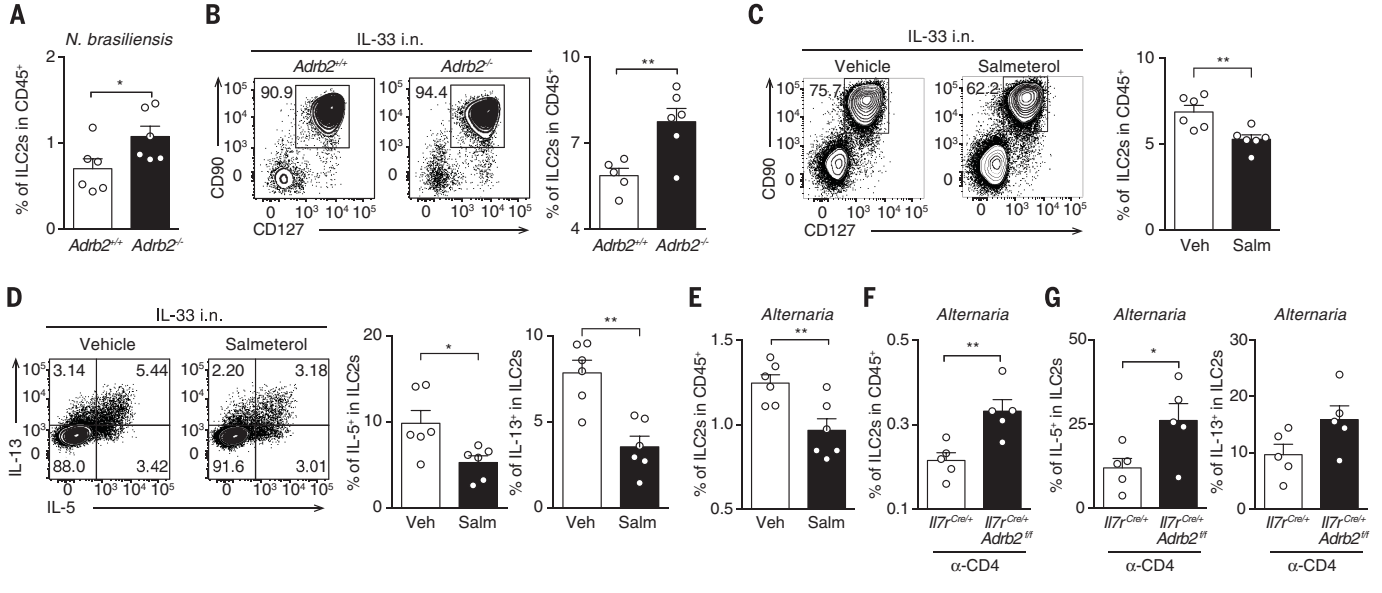


Fig. 3. β_2AR signaling inhibits ILC2 responses in lung inflammation. (**A**) Lung cells from *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ mice were analyzed 4 days after *N. brasiliensis* infection. Shown is an enumeration of CD127$^+$CD90$^+$Lin$^-$CD45$^+$ ILC2 percentages. (**B** to **D**) *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ mice (B) and B6 [(C) and (D)] mice treated with β_2AR agonist salmeterol (Salm) or vehicle were intranasally (i.n.) administered IL-33 for 3 days and analyzed 4 days later. Shown are flow cytometry plots of Lin$^-$CD45$^+$ cells [(B) and (C)] and CD127$^+$CD90$^+$Lin$^-$CD45$^+$ ILC2s (D). Also shown are enumerations of CD127$^+$CD90$^+$Lin$^-$CD45$^+$ ILC2 percentages [(B) and (C)] and cytokine production (D). The numbers in flow cytometry plots represent percentages in each gate. (**E**) B6 mice treated with β_2AR agonist salmeterol or vehicle were intranasally administered an *Alternaria* extract for 3 days and then analyzed 4 days later. Shown is an enumeration of CD127$^+$CD90$^+$Lin$^-$CD45$^+$ ILC2 percentages. (**F** and **G**) *Il7r*$^{cre/+}$ and *Il7r*$^{cre/+}$ *Adrb2*$^{f/f}$ mice treated with α-CD4 mAb were intranasally administered an *Alternaria* extract for 3 days and then analyzed 4 days later. Shown are enumerations of cytokine production. For all panels, each circle represents data from one mouse, bar graphs represent mean ± SEM, and data are representative of two experiments. *P < 0.05 and **P < 0.01 by unpaired two-tailed Student's t test.

significantly higher in *Adrb2*$^{-/-}$ mice than in *Adrb2*$^{+/+}$ mice (Fig. 4F). Together, these data indicate that β_2AR signaling negatively regulates ILC2 proliferation in vivo during type 2 inflammation.

To examine whether β_2AR-dependent inhibition of ILC2 proliferation is cell intrinsic, sorted ILC2s were stained with a proliferation dye (CellTrace Violet) and then cultured in vitro with a β_2AR agonist. Agonist-treated ILC2s exhibited significantly higher levels of CellTrace Violet intensity compared to vehicle-treated ILC2s, suggesting that the reduced proliferation after β_2AR agonist treatment occurs through a cell-intrinsic mechanism (Fig. 4G). To further test this

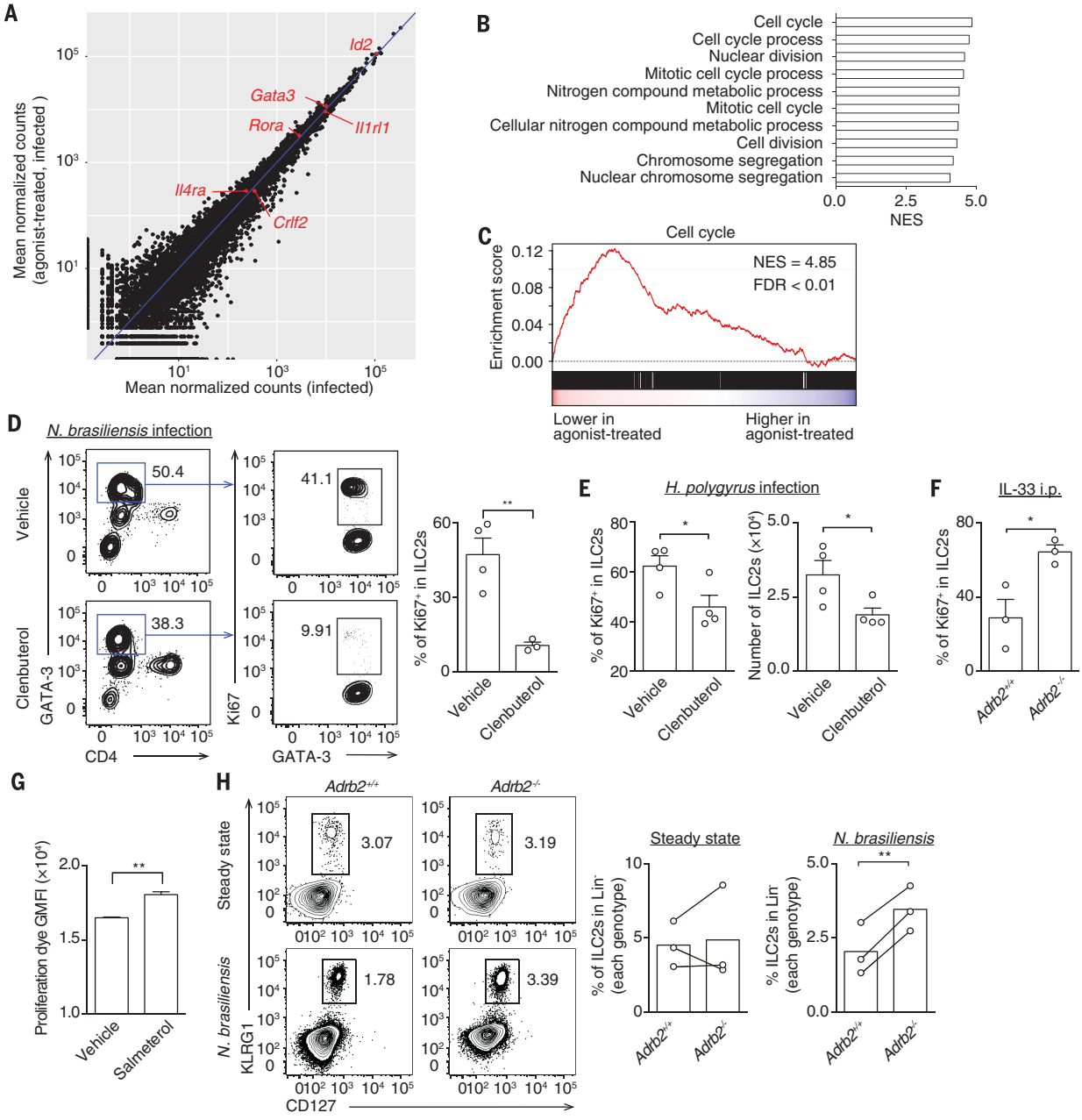


Fig. 4. β_2AR stimulation negatively regulates ILC2-intrinsic cell proliferation. (**A** to **C**) KLRG1$^+$CD127$^+$CD90$^+$Lin$^-$CD45$^+$ mLN ILC2s sorted from *N. brasiliensis*–infected *Adrb2*$^{+/+}$ mice treated with or without clenbuterol were analyzed by RNA-seq. Shown are (A) a scatter plot showing mean of normalized counts, (B) bar graphs of the significantly enriched [false discovery rate (FDR) < 0.01] GO terms with the highest normalized enrichment scores (NES), and (C) a GSEA plot of the GO term "cell cycle." (**D** and **E**) B6 mice treated with clenbuterol or vehicle were analyzed for percentages of Ki67$^+$ in ILC2s (D) 4 days after *N. brasiliensis* infection and (E) 8 days after *H. polygyrus* infection. Data are representative of two experiments. (**F**) *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ mice intraperitoneally (i.p.) administered IL-33 were analyzed for Ki67$^+$ in ILC2s. Data are representative of two experiments. (**G**) Sorted SI ILC2s stained with cell-proliferation dye were cultured with IL-2, IL-7, IL-33, and β_2AR agonist salmeterol or vehicle for 4 days, and the enumeration of proliferation dye geometric mean fluorescent intensity (GMFI) is shown. $n = 3$. Data are representative of three experiments. For panels (D) through (G), each circle represents data from one mouse, and bar graphs represent mean ± SEM; $^*P < 0.05$ and $^{**}P < 0.01$ by unpaired two-tailed Student's t test. (**H**) Chimeric mice reconstituted with congenic *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ BM cells were analyzed at steady state and 7 days after *N. brasiliensis* infection. Flow cytometry plots and enumerations of ILC2 percentages for each genotype are shown. Connected symbols represent data from one chimeric mouse. Data are representative of two experiments. $^{**}P < 0.01$ by paired two-tailed Student's t test.

in vivo, we reconstituted irradiated B6 mice with congenically marked *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ BM cells and compared the proliferation of *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ ILC2s in the same chimeric mouse (Fig. 4H). At steady state, the chimeric mice had similar frequencies of *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ ILC2s, in accordance with comparable ILC2 frequencies in naive *Adrb2*$^{+/+}$ and *Adrb2*$^{-/-}$ mice (fig. S2, B to D). However, after *N. brasiliensis* infection, the frequencies of *Adrb2*$^{-/-}$ ILC2s were increased compared to the frequencies of *Adrb2*$^{+/+}$ ILC2s (Fig. 4H), indicating that β_2AR signaling inhibits ILC2 proliferation in a cell-intrinsic manner in vivo.

Collectively, these results reveal a previously unrecognized regulatory circuit that operates between the adrenergic nervous system and the innate immune system to control type 2 inflammation at multiple mucosal sites. Specifically, signaling through β_2AR negatively regulates ILC2 proliferation and effector function. Given the importance of ILC2s in driving type 2 immune responses (*10*, *26*, *27*) and the altered β_2AR expression levels on ILC2s during inflammation, it appears that β_2AR may function as a molecular rheostat to fine-tune the ILC2 response, thus controlling the balance between promotion of host-protective acute ILC2 responses and prevention of chronic pathologic type 2 inflammation. Intriguingly, in contrast to the adrenergic-mediated negative regulation observed here, recent studies have identified cholinergic neurons as potent activators of intestinal and lung ILC2s through the production of the neuropeptide neuromedin U (*20*–*22*). Thus, the mammalian nervous system appears to have evolved dual mechanisms to rapidly activate or repress these innate immune cells to protect the host against diverse inflammatory stimuli.

REFERENCES AND NOTES

1. B. Pulendran, D. Artis, *Science* **337**, 431–435 (2012).
2. H. Hammad, B. N. Lambrecht, *Immunity* **43**, 29–40 (2015).
3. R. M. Locksley, *Cell* **140**, 777–783 (2010).
4. R. M. Anthony, L. I. Rutitzky, J. F. Urban Jr., M. J. Stadecker, W. C. Gause, *Nat. Rev. Immunol.* **7**, 975–987 (2007).
5. D. Artis, H. Spits, *Nature* **517**, 293–301 (2015).
6. A. B. Molofsky, A. K. Savage, R. M. Locksley, *Immunity* **42**, 1005–1019 (2015).
7. G. Eberl, M. Colonna, J. P. Di Santo, A. N. McKenzie, *Science* **348**, aaa6566 (2015).
8. K. Moro *et al.*, *Nature* **463**, 540–544 (2010).
9. D. R. Neill *et al.*, *Nature* **464**, 1367–1370 (2010).
10. A. E. Price *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11489–11494 (2010).
11. H. Morita, K. Moro, S. Koyasu, *J. Allergy Clin. Immunol.* **138**, 1253–1264 (2016).
12. D. L. Felten, S. Y. Felten, S. L. Carlson, J. A. Olschowka, S. Livnat, *J. Immunol.* **135** (suppl.), 755s–765s (1985).
13. S. Livnat, S. Y. Felten, S. L. Carlson, D. L. Bellinger, D. L. Felten, *J. Neuroimmunol.* **10**, 5–30 (1985).
14. J. B. Furness, *Nat. Rev. Gastroenterol. Hepatol.* **9**, 286–294 (2012).
15. E. M. Sternberg, *Nat. Rev. Immunol.* **6**, 318–328 (2006).
16. J. C. Nussbaum *et al.*, *Nature* **502**, 245–248 (2013).
17. S. Talbot *et al.*, *Neuron* **87**, 341–354 (2015).
18. L. Galle-Treger *et al.*, *Nat. Commun.* **7**, 13202 (2016).
19. I. J. Elenkov, R. L. Wilder, G. P. Chrousos, E. S. Vizi, *Pharmacol. Rev.* **52**, 595–638 (2000).
20. V. Cardoso *et al.*, *Nature* **549**, 277–281 (2017).
21. C. S. N. Klose *et al.*, *Nature* **549**, 282–286 (2017).
22. A. Wallrapp *et al.*, *Nature* **549**, 351–356 (2017).
23. S. Ibiza *et al.*, *Nature* **535**, 440–443 (2016).
24. E. C. Mackley *et al.*, *Nat. Commun.* **6**, 5862 (2015).
25. T. Hoyler *et al.*, *Immunity* **37**, 634–648 (2012).
26. S. H. Wong *et al.*, *Nat. Immunol.* **13**, 229–236 (2012).
27. C. J. Oliphant *et al.*, *Immunity* **41**, 283–295 (2014).
28. P. Panina-Bordignon *et al.*, *J. Clin. Invest.* **100**, 1513–1519 (1997).

ACKNOWLEDGMENTS

We thank the D. Artis lab members and the G. F. Sonnenberg lab members for discussion and critical reading of the manuscript. We also thank D. Mucida (The Rockefeller University) and P. A. Muller (The Rockefeller University) for advising on TH staining, S. Thomas (University of Pennsylvania), G. Karsenty (Columbia University), P. Liu (the Wellcome Trust Sanger Institute), and J. Sun (Memorial Sloan Kettering Cancer Center) for mice. We also thank D. Farber (Columbia University) and LiveOnNY for human lung samples. **Funding:** This work was supported by grants from The Naito Foundation (to S.M.), the Japan Society for the Promotion of Science (JSPS) Overseas Research Fellowships (to S.M.), the Novo Nordic Foundation (grant 14052 to J.B.M.), the German Research Foundation (grant KL 2963/1-1 to C.S.N.K), an Australian National Health and Medical Research Commission Early Career Fellowship (to L.C.R.), the Jill Roberts Institute (G.G.P.), a Weill Cornell Medicine Pre-Career Award (to L.A.M.), the NIH (grants F32-DK109630-01 to N.A.Y.; F32-AI134018-01 to L.A.M.; and AI061570, AI087990, AI074878, AI083480, AI095466, AI095608, AI102942, and AI097333 to D.A.), a European Research Council Advanced Grant (742883 to H.-R.R.), the Burroughs Wellcome Fund (grant to D.A.), and the Crohn's & Colitis Foundation of America (grant to D.A.). **Author contributions:** S.M., J.R.B., A.-L.F., J.B.M., C.S.N.K., L.C.R., L.A.M., and D.A. designed and performed the research. S.M., C.S.N.K., G.G.P., N.A.Y., and D.A. analyzed the data. H.-R.R. provided the *Il7r*^{cre} mice. S.M. and D.A. wrote the manuscript with input from the other authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data presented in this manuscript are tabulated in the main paper and in the supplementary materials. RNA-seq data are deposited under accession number GSE108884 in the Gene Expression Omnibus database.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6379/1056/suppl/DC1
Materials and Methods
Figs. S1 to S3
Table S1
References (*30*–*45*)

30 April 2017; resubmitted 15 November 2017
Accepted 18 January 2018
10.1126/science.aan4829

Apply for our exciting research Prize!

SCIENCE-PINS PRIZE FOR NEUROMODULATION

SCIENCE-PINS

$25, 000 Grand Prize
Get published in *Science*!

The *Science*-PINS Prize is a new, highly competitive international prize that honors scientists for their excellent contributions to neuromodulation research. For the purposes of the Prize, neuromodulation is any form of alteration of nerve activity through the delivery of physical (electrical, magnetic, or optical) stimulation to targeted sites of the nervous system.

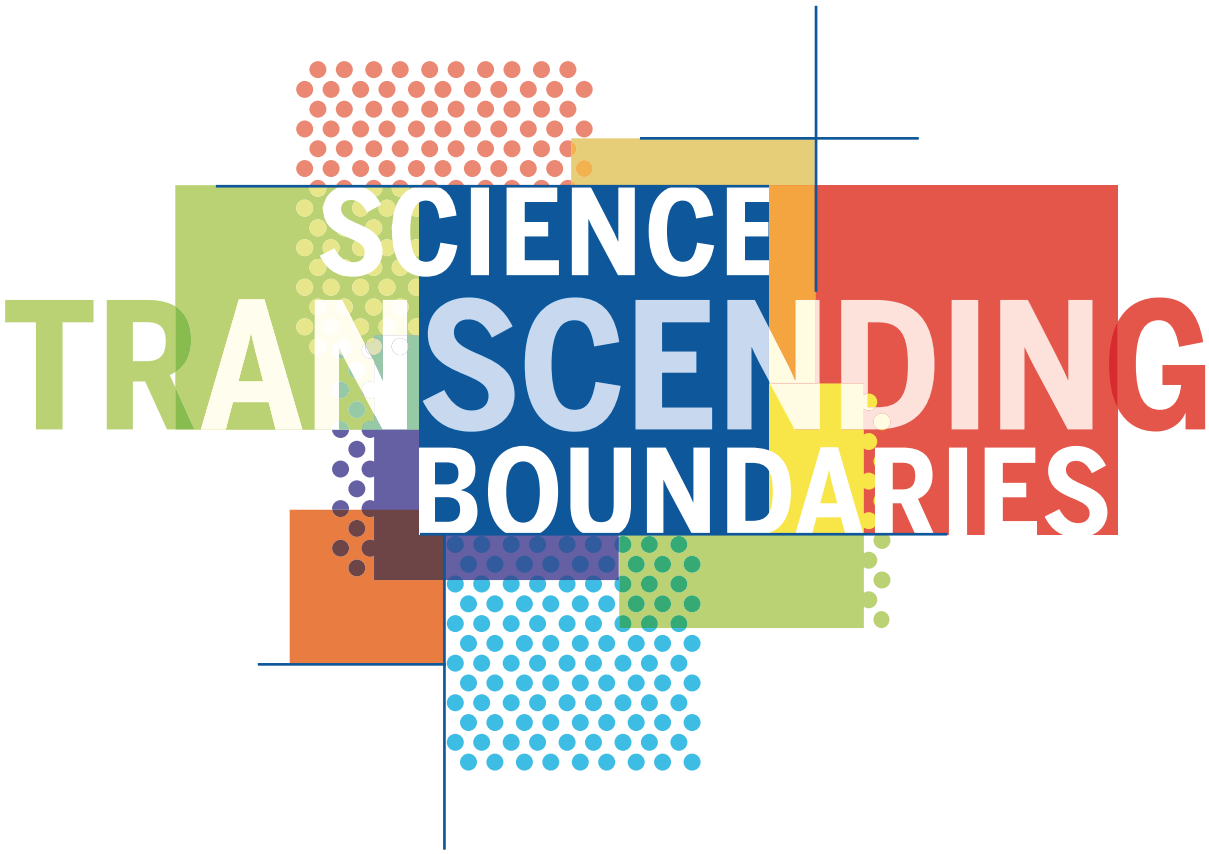
For full details, judging criteria, and eligibility requirements, visit:

www.sciencemag.org/prizes/pins

Submission Deadline: March 15, 2018

Science AAAS

PINS 品驰



Call for Session Proposals

Session proposals for the 2019 AAAS Annual Meeting are now being solicited. Deadline for submission: **April 19, 2018**

aaas.org/meetings

AAAS
ANNUAL MEETING
Feb. 14–18, 2019 | Washington, DC

How can science help address the many divisions in our communities, in global society, and in science itself? Science plays a unique and important role in how people see and understand the world, and how lines and distinctions are drawn. At this Annual Meeting, we look for ways science is bringing together people, ideas, and solutions from across real and artificial borders, disciplines, sectors, ideologies, and traditions. How can science working across boundaries improve its ability to find solutions to the pressing problems of our age? How can scientists, wherever they work, more effectively engage with the broader society? How can we find better ways to engage the public, especially in expanding access to science and scientific careers? At the international level, science diplomacy builds bridges between countries. How can we encourage more of this, and utilize science as a common ground more locally as well? Can science contribute information that might reduce or mitigate the starkly divergent interests of different populations and demographics, such as urban and rural communities? What boundaries most impede your research or your career? While acknowledging that some boundaries are useful and necessary, the meeting theme considers how research can be applied to problematic separations in the world, and how unhelpful boundaries within science are being addressed.

NEW ENGLAND BioLabs® Inc.

be INSPIRED
drive DISCOVERY
stay GENUINE

Did you get the message?

In recent years, the discovery of new classes and modifications of RNA has ushered in a renaissance of RNA-focused research. Did you know that NEB® offers a broad portfolio of reagents for the purification, quantitation, detection, synthesis and manipulation of RNA? Experience improved performance and increased yields, enabled by our expertise in enzymology.

FEATURED PRODUCTS *for* RNA RESEARCH:

NEW Monarch® Total RNA Miniprep Kit – rapidly purify up to 100 µg of high quality total RNA from multiple sample types

Luna® One-Step RT-qPCR kits – improve RT-qPCR performance with our novel, Luna WarmStart® Reverse Transcriptase

NEW NEBNext® Ultra™ II RNA Library Prep kits – generate high quality RNA-seq libraries, even with limited amounts of RNA

HiScribe® *in vitro* transcription kits – rapidly synthesize high yields of high quality RNA

Let NEB help streamline your **RNA-related workflows**.
Get started at **NEBrna.com**

Waseda University pushes forward with global academic network

Innovative programs and prioritized funding have propelled Waseda University to record highs in world university rankings, underscoring the university's reputation for openness, dynamism, and diversity.


Shuji Hashimoto

Located in Tokyo, Waseda University is one of Japan's highest-ranked private universities, according to independent surveys published by the 2017 Quacquarelli Symonds (QS) World University Rankings. "With over 7,000 international students enrolled throughout the year and nearly 800 exchange agreements, and with overseas partners in 91 countries, Waseda is Japan's top destination for highly motivated international students," says **Shuji Hashimoto, senior executive vice president for academic affairs and university provost.** "Importantly, more than 4,000 students at Waseda travel overseas each year; this is unprecedented in Japan."

This global academic network is an important reason why Waseda University was selected in September 2014 for the highly competitive, multimillion dollar, 10-year, Top Global University (TGU) Project by Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT). Funding from the TGU Project as well as the university's own financial support is being targeted for seven key educational and research areas to further enhance Waseda's international standing: (1) Frontier of Embodiment Informatics: Information Communications Technology (ICT) and Robotics; (2) Energy and Nanomaterials; (3) Global Japanese Studies; (4) Multiscale Analysis, Modelling, and Simulation; (5) Positive/Empirical Analysis of Political Economy; (6) Health Promotion: The Joy of Sports and Exercise; and (7) Global Asia Studies. "With over 50,000 students, 5,500 academic staff, and 7 undergraduate and 13 English-based graduate programs, it was imperative to prioritize into key units," says Hashimoto.

The strategy is yielding very favorable results, with 2017 QS World University Rankings data showing Waseda University ranked sixth in Japan overall, and first in sports-related subjects. In world rankings, the university has nine fields of research in the top 100.

"We attribute these results to the international collaboration of our key units and the recruitment of overseas faculty members through the TGU Project," says Hashimoto. "We are on the right track to achieve our goal of creating and sharing knowledge for the betterment of society on a global scale."

Below and in upcoming articles, the coordinators of the seven key units at Waseda University describe their achievements and plans for the future.

Waseda University
www.waseda.jp/top/en

Coexistence of robots and humans —Frontier of Embodiment Informatics: ICT and Robotics

Waseda University created the world's first humanoid in 1973—WABOT-1 (WAseda roBOT-1)—by fusing ideas from mechanical and electrical engineering. Now, **Shigeki Sugano, head of the Frontier of Embodiment Informatics: Information and Communications Technology (ICT) and Robotics Unit,** is leading Waseda's research on robotics by collaborating with researchers in Germany and Singapore. "The day will come when a robot will say, 'Please get out of my way,' to humans it encounters as it moves along a walkway," says Sugano. "Just as cars and humans coexist, robots and humans will share the same space. But we still have many technical and social issues to resolve."

Such symbiosis between robots and humans requires robots to adapt to the unpredictable behavior of humans. Sugano is developing robots with ultraflexible joints to enable them to perform complicated and dynamic movements; for example, rapid twisting and turning actions of robotic arms are required to avoid collisions with people on a busy street.

Recently, Sugano and his group reported on "training" flexible joints of robots to perform complex dynamic tasks using deep learning,[1] which allows a reduction in the number of training iterations required.

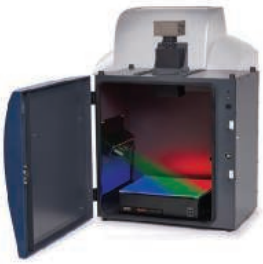

Shigeki Sugano

But despite growing demand for robots, Sugano highlights their mass production as a major issue for their proliferation. "Manufacturing intelligent robots is very complicated," says Sugano. "There is not one single axis for robot development. They are like smartphones, but can move!"

Training the next generation of robotics engineers and scientists is Sugano's primary goal. "We must nurture experts by establishing new educational programs with innovative, interdisciplinary curricula and strong interaction with industry."

[1]K. Takahashi, T. Ogata, J. Nakanishi, G. Cheng, S. Sugano, *Adv. Rob.* **31**, 1002–1015 (2017), doi: https://doi.org/10.1080/01691864.2017.1383939.

Frontier of Embodiment Informatics: ICT and Robotics Unit
www.waseda.jp/inst/sgu/en/unit/ict-robotics



Chemiluminescence Imaging

The GeneGnome XRQ is dedicated to chemiluminescence imaging. This system is built for high performance and automation, and now features a next-generation, high quantum efficiency Syngene CCD camera for even greater sensitivity. It automatically selects the right imaging conditions for any Western blot, irrespective of the reagents being used, and can easily handle all chemiluminescence applications, producing superb images even from the faintest of signals. The compact instrument sits on any laboratory bench, taking up very little space. It is available in three variants: GeneGnome XRQ with monitor and built-in processor, GeneGnome XRQ NM with built-in processor but no monitor (use your own monitor), and GeneGnome XRQ NPC with no monitor or built-in processor (use your own monitor and computer). The system comes complete with unlimited copies of GeneTools analysis software.

Syngene
For info: 800-686-4407
www.syngene.com/genegnome-xrqa

3D Cell Culture Platform

Mimetix scaffolds mimic the extracellular matrix by providing an ideal architectural environment to support the growth of cells in 3D. They are created by electrospinning the medical-grade polymer poly(L-lactide) into microfibers, which are highly consistent with regard to fiber diameter and pore size, resulting in excellent reproducibility of cell-based assays. The scaffolds are incorporated into standard SBS-footprint well-plate frames with bases of superior optical clarity and minimal base distortion. The 50-µm scaffold depth is thick enough to provide the benefits of 3D cell morphology and behavior, yet thin enough to allow microscopic imaging.

AMS Biotechnology
For info: +44-(0)-1235-828-200
www.amsbio.com/mimetix.aspx

Human Platelet Lysate

PLTMax Human Platelet Lysate is a nonxenogeneic, animal serum–free product derived from human platelets. It is a growth factor–rich supplement and a superior alternative to fetal bovine serum for human mesenchymal stem cell culture and other primary cells. PLTMax is used as a manufacturing component in the generation of adult stem cells in clinical trials in North America, Europe, South America, the Middle East, Asia, and Australia (Phase I to Phase III) in indications including neurology, nephrology, gastrointestinal disease, wound repair, and cardiology.

Biological Industries
For info: 972-(0)-4-996-0595
www.bioind.com/pltmax-human-platelet-lysate

Laboratory-Animal Serology Products

XpressBio manufactures and markets laboratory-animal serology products and services to the bioscience research community around the world. We specialize in ELISA immunoassays and reagents for mice, rats, guinea pigs, hamsters, rabbits, canines, felines, and nonhuman primates. In addition, we offer customizable solutions for your research needs, including PCR controls and assays; *Helicobacter* PCR genus and six species; mouse, rabbit, and human vaccine research ELISA kits; HIV p24 ELISA kits, including extended range; HIV integrase assays; HCV core ELISA kits; DNA purification reagents; and cytokine assays. Our Health Monitoring ELISA products range from an array of assays such as MHV, MPV, PVM, RPV, ECUN, MAD, RHDV, to a very unique native mouse parvo ELISA kit, mouse and rat *Pasteurella*, and *Helicobacter* ELISA kits. Our customizable ELISA assays are unique in the industry. We also manufacture toxicity ELISA assays and vaccine research assays.

XpressBio
For info: 301-228-2444
www.xpressbio.com

Water Baths

Thermo Scientific Precision water baths are setting a new standard with enhanced features that simplify workflows and maximize productivity. Their rugged construction, compact footprint, advanced temperature controls, ease of use, and safety features clearly make these water baths a smart choice for your smart lab. Choose from a variety of models—general purpose, circulating, coliform, shaking, and Dubnoff. Save valuable benchtop space with compact footprints, and protect your work with safety features including audible alarms and adjustable digital and overtemperature protection. Increase productivity with new auto-on and auto-off timers and an easy icon-based user interface. Access worldwide service and support when you need it, for added peace of mind.

Thermo Fisher Scientific
For info: 800-955-6288
www.thermofisher.com

SILAC Mouse Feed

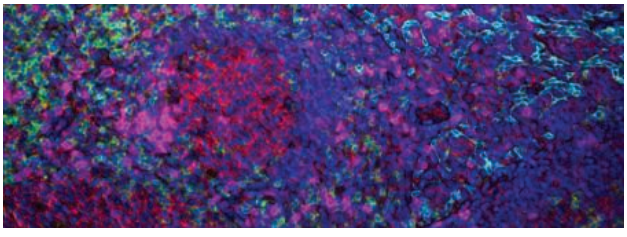
Cambridge Isotope Laboratories offers Mouse Express neutron coding (NeuCode) mouse feed with various combinations of labeled lysine isotopologues for high-resolution mass spectrometry proteomics studies. The mouse is the most commonly used animal model in biological and biomedical research. A promising new labeling technique for increasing the multiplexing potential of traditional SILAC (stable isotope labeling with amino acids in cell culture) quantitative experiments involves NeuCode. The ability to incorporate NeuCode-labeled media into the proteomes of actively growing mouse cells is beneficial for molecular phenotyping and screening studies. Kits are available in 1-week or 3-week quantities (each kit contains enough feed for three mice) and are made to order and always fresh.

Cambridge Isotope Laboratories
For info: 800-322-1174
www.isotope.com

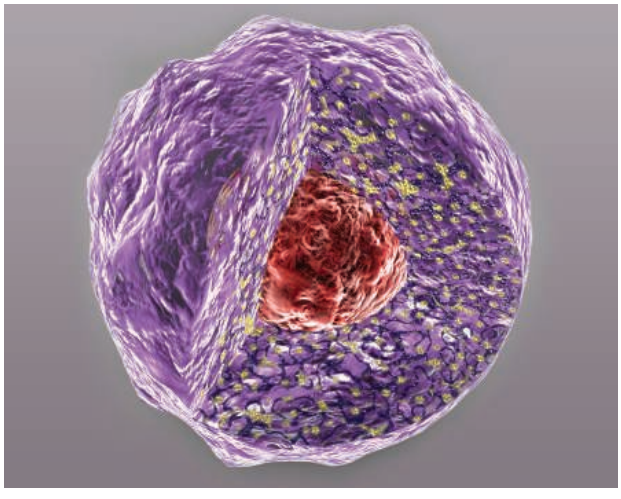
Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

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.

want new technologies?



antibodies
apoptosis
biomarkers
cancer
cytometry
data
diseases
DNA
epigenetics
genomics
immunotherapies
medicine
microbiomics
microfluidics
microscopy
neuroscience
proteomics
sequencing
toxicology
transcriptomics



watch our webinars

Learn about the latest break-throughs, new technologies, and ground-breaking research in a variety of fields. Our expert speakers explain their quality research to you and answer questions submitted by live viewers.

VIEW NOW!
webinar.sciencemag.org

Science
AAAS

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

@SciMagWebinars


RESEARCH ON THE EDGE
43.1355°N | 70.9443°W
COLLEGE WOODS, DURHAM, NH
OUR ECOLOGY RESEARCH IS SECOND TO ONE.
At the University of New Hampshire, understanding our natural world is second nature. Because of our commitment to researching and protecting the fields, forests, streams and oceans that surround us, a recent study in Ecosphere placed UNH second in North America in scholarly productivity in the field of ecology.
Welcome to the edge of possible — where intellectual curiosity is rewarded with exciting research opportunities and challenging academic programs. Whether examining the microbes that live in soils or tracking atmospheric gases high above the Arctic, we exist to help our students create possibilities for our world.
unh.edu/edge
University of New Hampshire

Science Career Fair
at Harvard University

March 14, 2018
11:00 AM – 5:00 PM EST

A Day of Job Opportunities and Career Seminars

JOB SEEKERS! ***Science*** **Careers** has partnered with Harvard University to produce a unique career fair open to all. Join us for a chance to meet with top scientific organizations and get important advice from career experts. The combination of valuable career development seminars, company presentations, and exciting career opportunities makes this a free "must-attend" event for scientists.

Northwest Science Building, Harvard Campus
B100
Cambridge, MA

For more details and to register, visit
sciencemag.org/careers/jobfair

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

SCIENCECAREERS.ORG

MUSC
MEDICAL UNIVERSITY of SOUTH CAROLINA

Recruiting Faculty in Biochemistry and Molecular Biology

Faculty positions at the Assistant and Associate Professor ranks are available in the Department of Biochemistry and Molecular Biology at the Medical University of South Carolina (MUSC) in Charleston. Candidates with research interests in all areas of biochemistry, cell, molecular, and cancer biology are welcome to apply. State-of-the-art laboratories, outstanding resources and research support are available.

Candidates for the Assistant and Associate Professor ranks should have a national reputation and a solid record of collaborative and peer-reviewed funded research. We are seeking outstanding scientists who would complement and expand existing research foci and programs at MUSC.

Located on the Atlantic coast in South Carolina, Charleston boasts one of the nation's most historic downtown areas, beaches and year-round outdoor life, as well as international cultural events such as the Spoleto Festival USA. (https://charlestoncommunityguide.com)

Interested researchers should send their CV and a summary of future research plans to:

Philip H. Howe, Ph.D.
Professor & Chair
Department of Biochemistry & Molecular Biology
Medical University of South Carolina
173 Ashley Avenue, MSC 509
Charleston, SC 29425
albano@musc.edu

MUSC is an Equal Opportunity Employer, promoting workplace diversity.

AWARDS

WIN Rising Star Award in Nanoscience and Nanotechnology

~ Call for Submissions ~

The Waterloo Institute for Nanotechnology (WIN) at the University of Waterloo is pleased to announce the call for submission for the WIN Rising Star Award in Nanoscience and Nanotechnology. The award recipient will receive $10,000 (CAD) as an honorarium at the International Symposium on Frontiers of Nanoscience and Nanotechnology on June 6, 2018.

Eligibility:
A full-time faculty member who has received their PhD in science or engineering in 2008 or later. The individual's research interest must align with one or more of the thematic area(s) of WIN: Smart and Functional Materials; Connected Devices; Next Generation Energy Systems; and Therapeutics and Theranostics. Details about these areas are available at www.nano.uwaterloo.ca.

Award:
Along with the honorarium for $10,000 (CAD), the award recipient will present a featured research seminar at the International Symposium. The award recipient will spend two days at WIN interacting with leading faculty members and graduate students.

Submission:
Deadline for submissions is **April 15, 2018 at 17:00**, via email to **WIN-RS@uwaterloo.ca**. Please visit the WIN website for more details: **www.nano.uwaterloo.ca**.

UNIVERSITY OF WATERLOO | WATERLOO INSTITUTE FOR nanotechnology

Genome editing

Immunology

Neutron scattering

Performance aware programming

Public health

Ultracold atom

Croucher Summer Courses 2018

A Croucher Summer Course is a residential course held over five or six days in Hong Kong. Participants have an opportunity to learn from world-renowned scientists, network with peers, participate in engaging and interactive discussions and, depending on the subject of the course, gain practical experience in the laboratory.

Each Croucher Summer Course is organised in partnership with a university in Hong Kong.

Registration fees cover course materials, accommodation, and meals. Places are limited; early application is advised:
www.croucher.org.hk

The Croucher Foundation is an independent private foundation dedicated to promoting the standard of the natural sciences, technology and medicine in Hong Kong.

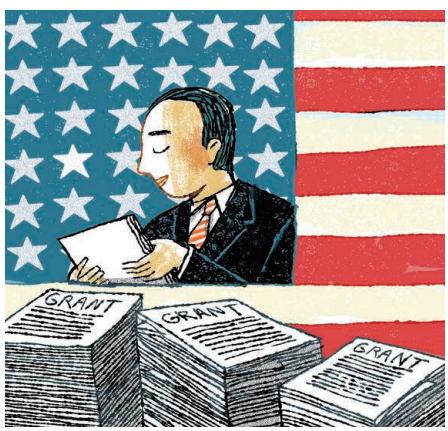
Croucher Foundation
裘槎基金會

By **Prosanta Chakrabarty**

My year as a fed

A friend had warned me that the entire Washington, D.C., region is an ARE (Acronym Rich Environment). On the first day of my 1-year rotation as an NSF (National Science Foundation) program director, I was finding out just how true that was. After the NEW (New Employee Workshop), I felt my eyes glaze over (perhaps blurred by tears) as I scanned several pages of "Common Acronyms of NSF." What had I gotten myself into?

In my regular life as a professor and ichthyologist, I discovered new species in far-off tropical countries and taught classes to enraptured masses. (OK, maybe indifferent flocks, but it was fulfilling nonetheless.) Now, I was about to leave that behind to spend a year allocating grant funding to other people. Did I really want to learn acronyms and push papers for the federal government? I reminded myself what I had said during my interview: I wanted to give back to NSF, which had been so good to me; I wanted to learn how the system worked; and I wanted to see my field from the top.



"I've returned to my academic work with a new, reinvigorated perspective."

My rotation ended a few months ago, and I'm now happily back at my university. But I wouldn't say that I'm back where I started. I've returned to my academic work with a new, reinvigorated perspective on what it means to be a scientist, how our government funding agencies support us, and why we need to support them. Despite the challenge of learning fed-speak and the seemingly endless training, I found that I actually liked learning how the proverbial sausage was made—even though it wasn't always pretty.

I gained the greatest insights from running grant review panels made up of scientists we brought in to help determine which proposals most deserved support. These panels were intellectually stimulating, exhilarating, and exhausting. In the long hours we spent debating the merits, methods, and minutiae of each proposal, I learned a great deal about how scientists read, understand, and evaluate the work of others. One epiphany was that many of the most successful grant writers were those who could not only explain what they were going to do and why, but could also get a panelist excited enough to advocate for their work. I realize now that *everything* I do should have a "hook" to reel people in. To convince others that a project is interesting and important, I need to articulate why I am excited about it.

Once we picked the top proposals, I had to get them funded. That forced me into a new role of advocating for someone else's science. I needed to justify all of my funding decisions to higher-ups so that those higher-ups could defend their program decisions to people even higher up, and so on. I valued being part of this long unbroken chain that ultimately explains why it is worth supporting cutting-edge science to the nonscientists overseeing federal funding: Congress and the president. I found the process inspiring, even if I was sometimes a little jealous of the superb proposals I found myself promoting.

Although the ARE at NSF was hard to get used to at first, I'm glad I ventured outside my comfort zone. Now I know how much work goes into not only doing good science, but also funding it. Coming back to my home institution, I realized that my science advocacy shouldn't end with my federal employment. Just as it was my role as a rotator to be the voice of the community inside the walls of NSF, on the outside I can pull back the curtain a bit for those who want to know more about the granting process by giving seminars or one-on-one advice. I have also become an advocate for supporting funding agencies more generously so that they can support deserving scientists. I never thought I'd be holding up signs at a march defending science or going on the offensive protecting funding agencies using social media, but I can't shake my newfound understanding of the need to promote the interests of science. By bridging the knowledge gaps between government and academia, we can ensure that scientists get the support we need. ■

Prosanta Chakrabarty is an associate professor and curator at Louisiana State University in Baton Rouge.