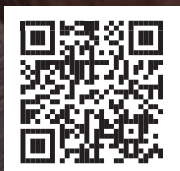# Science

PANDEMICS IN WAITING

Hunting for emerging diseases in the Amazon p. 450

NEWS FROM Science

Up-to-the-minute research and policy news you won't find in print

Visit us online to read all the news coverage that
there just wasn't enough room to print in this issue.

ScienceMag.org/news

CONTENTS

In Manaus, Brazil, a city confronts the rainforest and its pathogens.

450

534

ON THE COVER

In a fragment of Amazon rainforest within the city of Manaus, Brazil, parasitologist Aline Ramos holds samples of feces, bodily fluids, and tissue collected from pied tamarin monkeys. Ramos is part of a team screening rainforest animals for novel and known viruses and parasites that might spill over into humans, possibly sparking an outbreak. See page 450. *Photo: Hilaea Media/Dado Galdieri*

Science ▲AAAS

Equity in cost-benefit analysis

One key way that many governments around the world incorporate scientific research into policy-making is through cost-benefit analysis (CBA). But despite well-established practices for rigorous estimation of the pros and cons of policies, there is room to improve, particularly in characterizing difficult-to-measure benefits and the distribution of the costs and benefits across different segments of society. In this regard, announcements by US President Biden, if brought to fruition, could have far-reaching implications for how CBA is used in government decision-making. But such promising (and familiar) rhetoric is no guarantee of real progress, and the devil, as always, is in the details. These details are where the scientific community has an important role to play to improve the use of CBA and to hold the administration accountable.

Some 40 years ago, President Reagan ushered in a regulatory "revolution" by requiring CBA for major regulations. Many other countries have since followed suit. Different US presidents have emphasized different parts of CBA. Republican presidents have tended to focus on the cost of regulation, or cost-benefit balancing. Democratic presidents have tended to place increasing emphasis on benefits and distribution. A Biden memo in January revisits many of these traditional Democratic priorities, but recognizing heightened awareness in society and politics, expands benefits and equity further. The memo could be construed as asking for a roadmap on how to address equity concerns that arise.

It is useful to divide this problem into three parts: getting better information on the likely winners and losers from a policy; examining how to include that information in a CBA; and specifying how the distributional impacts should enter into final regulatory decisions. Better information on the distributional aspects of regulation could be very useful but will not be free. If the administration is interested in learning more about equity, it should consider funding research. This can be broken down into two parts: research on benefits and research on costs by, say, socioeconomic grouping. On the benefit side, this could involve developing better information on "dose" and response by groups in the case of environmental regulations. Groups could be defined by income or race, for

example. On the cost side, it would be desirable to know more about which groups bear the most substantial costs of regulation. This is relatively straightforward in some cases where costs are highly concentrated (e.g., in the case of plant closures) but is more challenging when trying to estimate the impact of a regulation on wages or prices across different groups.

Interestingly, almost no CBAs of regulations provide quantitative assessments of the benefits and costs of a regulation by particular subgroups in the population. One reason is that it is challenging to develop such estimates. Yet, such calculations are critical for understanding the distributional impacts of a regulation. Scientists can play an important role in defining how best to develop such estimates and the likely costs and benefits of doing so.

> "How…should… information on equity be used in the regulatory decision process?"

The current approach to factoring information on equity into a CBA is generally to exclude such information from the quantitative analysis altogether and only consider it in the final decision to regulate. One alternative is to introduce distributional weights in comparing benefits and costs across different socioeconomic groups. The precise formulation of these weights is likely to be highly contentious. Scientists may not have any higher authority in the determination of these weights, which is largely a question of adjudicating conflicts among competing societal values, but scientists can play an important role in calculating how different weights could affect the estimation of net benefits

How then should the additional information on equity be used in the regulatory decision process? One approach would be to ask the decision-maker to try to maximize net benefits but leave latitude to consider a wide range of factors, including equity. At the same time, the basis for reaching the decision, and the role that quantitative analysis played in reaching that decision, should be stated clearly. Scientists have an important role to play in estimating how decision rules are likely to affect overall net benefits and the net benefits going to different groups.

The question ahead for academics and practitioners is how best to address equity concerns if the goal is to promote a regulatory agenda that actually does more good than harm.

–Robert W. Hahn

Robert W. Hahn is a distinguished senior fellow at the Technology Policy Institute in Washington, DC, USA and an honorary research associate at the Smith School of Enterprise and the Environment at Oxford University, Oxford, UK. robert.hahn@smithschool.ox.ac.uk

IN BRIEF

Edited by
Sofia Moutinho

CLIMATE CHANGE
Biden increases U.S. climate pledge

The United States submitted its new goals to the Paris climate agreement last week, pledging to cut its planet-warming greenhouse gas emissions 50% to 52% below 2005 levels by 2030. The nonbinding pledge, made on Earth Day at an online climate summit convened by President Joe Biden, is one of the most aggressive targets of any wealthy country. The cuts are also higher than the 26% to 28% reduction by 2025 pledged by former President Barack Obama when the United States first joined the agreement in 2015. In the past month, Canada, Japan, and the United Kingdom have also committed to cuts steeper than their earlier pledges; other large polluters, such as China, India, and Russia, have yet to increase their goals ahead of a critical U.N. climate meeting this winter in Glasgow, U.K. The U.S. pledge has a long road to reality. It will require immediate increases in renewable energy, widespread adoption of electric vehicles, and other steps, many of which will require laws that could be difficult to pass, such as Biden's climate-focused infrastructure bill.

Vaccine injury payouts saved

VACCINATION | The Department of Health and Human Services (HHS) last week killed a rule that would have made it much more difficult for people who sustain shoulder injuries during vaccination to win compensation from a $4.1 billion government fund. The rule had been finalized on 19 January, the last day of former President Donald Trump's administration. But President Joe Biden's administration froze its implementation on 20 January. The new administration said the previous one had been "irregular in its haste" when it moved to remove shoulder injuries from a list of injuries in which the petitioner does not have to prove a vaccine caused the injury, making it easier to win a government payout (*Science*, 10 April 2020, p. 121).

Shoulder injuries accounted for nearly 55% of more than 2400 claims filed with the National Vaccine Injury Compensation Program in the past 2 years, most of them after flu shots. COVID-19 vaccines fall under a different HHS program, but if they win formal approval by the U.S. Food and Drug Administration, they could in principle be added to the national compensation program.

PHOTO: WHITE HOUSE

Journals to outsource peer review

PUBLISHING | Sixteen journals, including *BMJ Open Science* and *Royal Society Open Science*, say they will accept articles reviewed by the nonprofit Peer Community In Registered Reports (PCI RR). The organization, launched last week, will review one type of article: "registered reports," describing studies for which detailed experimental plans are peer reviewed before research begins. Once the research is complete, PCI RR will do a second round of peer review including results and analysis. Papers it recommends can then be published in any of the 16 journals without further review as long as they meet a journal's normal criteria. The organization, funded by donations, will provide peer review free to authors and journals in any discipline.

NIH seeks old drugs for COVID-19

COVID-19 | The National Institutes of Health (NIH) last week announced it would launch a large study repurposing existing drugs for patients with mild COVID-19 symptoms who don't need hospitalization. The $155 million trial, which aims to recruit 13,500 participants, will open in a few weeks at several research centers. It will test up to seven medications already approved for other conditions, but NIH hasn't yet named them. A recent, similar trial at the University of Oxford that has so far enrolled about 4800 patients has found that the asthma drug budesonide lessens symptoms and speeds recovery in certain COVID-19 patients.

COVID-19 raises pregnancy risks

MEDICINE | A large study has firmed up earlier evidence that SARS-CoV-2 increases the rate of complications for pregnant women and their babies. The study followed 706 pregnant women with COVID-19 and 1424 uninfected pregnant women at hospitals in 18 countries. Infected women had a 76% higher risk of developing problems caused by pregnancy-associated high blood pressure and a 59% higher risk of preterm birth. They were also five times more likely to be admitted to intensive care than uninfected women. Eleven women with COVID-19 died, compared with one uninfected woman, researchers report in *JAMA Pediatrics*. Infected women with fever and shortness of breath had babies with a fivefold increased risk of complications such as immature lungs, eye disorders, and brain damage. The coronavirus may affect pregnancy via changes in a woman's heart, lungs, and immune system.

IN FOCUS Large brown seaweeds known as laminarian kelp make up the most extensive coastal marine biome, a new analysis reveals. Last year, researchers ranked seagrass as the top underwater biome at 1.65 million square kilometers (km²), followed by kelp, coral reefs, and mangroves. New data from the Arctic expand global kelp forest estimates to more than 2 million km², or about 36% of coastal waters, as reported online on 20 April in *Biological Conservation*.

The results show pregnant women should be among priority groups for COVID-19 vaccines, the authors say. A separate study last week found no obvious safety problems in more than 800 U.S. women who gave birth after receiving messenger RNA vaccines.

Dawkins loses award over tweets

COMMUNITY | The American Humanist Association has decided to withdraw Richard Dawkins's 1996 Humanist of the Year award as a result of his "history of making statements that use the guise of scientific discourse to demean marginalized groups." The decision came soon after the evolutionary biologist and former University of Oxford professor wrote a tweet comparing transgender people to Rachel Dolezal, a civil rights activist who for years posed as Black. The association said the tweet "implies that the identities of transgender individuals are fraudulent, while also simultaneously attacking Black identity as one that can be assumed when convenient." In 2015, Dawkins argued that trans women are not women based on their chromosomes, but said he would use the pronoun "she" out of courtesy.

Cellphones detect earthquakes

SEISMOLOGY | The first large-scale, phone-based earthquake early warning system will be deployed in Greece and New Zealand, Google announced this week. Since last year, the company has been testing the use of data compiled from its more than 2 billion active Android phones to pinpoint the location and strength of earthquakes. The measurement comes from the built-in Android phones' accelerometers, which sense movement just like seismometers. When the phones detect earthquakelike signals, they alert a server that combines information from many phones. If enough phones corroborate the result, an alert goes out. The cellphone results compared well to those from

seismometer-based warning systems in Japan and the United States. Google chose New Zealand and Greece, both countries with high earthquake hazards and many Android phones, to premier the system because they lack operational warning systems of their own. Eventually, phone-based alerts could be available worldwide.

First U.S. test for GM mosquitoes

BIOTECHNOLOGY | Genetically modified mosquitoes designed to prevent the spread of viruses such as Zika and dengue are set to be released in the United States for the first time. Starting this week, the company Oxitec will free fewer than 12,000 transgenic *Aedes aegypti* mosquitoes, all of them nonbiting males, in the Florida Keys as part of a pilot study, the company announced on 23 April. They are engineered to carry a gene that kills their female offspring, reducing the population of mosquitoes capable of transmitting diseases. Field tests outside the United States have shown dramatic population drops, though the company has not published definitive evidence that the strategy reduces disease in humans. The project, which has long faced public opposition in Florida, won approval in May

Oxitec aims to reduce wild populations of *Aedes aegypti* mosquitoes that transmit viruses to humans.

2020 from the Environmental Protection Agency, which predicted no adverse effects on people or local wildlife.

Biden fills out science team

SCIENCE POLICY | President Joe Biden last week picked two veterans of government service and a newcomer to fill top science positions. He named soil scientist Asmeret Asefaw Berhe of the University of California, Merced, to lead the Department of Energy's Office of Science. Berhe, born in Eritrea, has little government experience but has won accolades for her research and efforts to promote diversity in science. She would become the first Black woman to lead the science office if confirmed by the Senate. Oceanographer Rick Spinrad of Oregon State University, Corvallis, who has held numerous posts at the National Oceanic and Atmospheric Administration, is Biden's choice to lead that agency. To run the State Department's science bureau, he chose Monica Medina, an ocean policy expert and attorney at Georgetown University. Confirmation hearings for the three nominees could come as early as next month.

Europe pitches strict AI rules

REGULATION | In one of the broadest attempts to regulate artificial intelligence (AI) to date, the European Commission on 21 April proposed new rules for algorithms that power everything from medical device and credit scoring software to chatbots and facial recognition systems. The rules divide AI technologies into risk categories, and put outright bans on some, such as systems that would score individuals' "social credit." Other "high-risk" systems, including those that collect biometric data, would require a strict vetting process. However, there are exceptions for national security, and it could take years before the rules become law: They must first pass the European Council and the European Parliament and be adopted by member countries.

National Academy ups diversity

COMMUNITY | An effort to increase gender, racial, and geographic diversity in the U.S. National Academy of Sciences (NAS) has begun to bear fruit. Nearly half of the 120-member 2021 class announced last week are women, compared with one-quarter in 2011. The new cohort includes nine Black scientists; NAS officials say previous classes never had more than three and often had none. "We need to do better, but I'm amazed at how far we've come," says plant geneticist Susan Wessler, NAS home secretary. The academy's governing council has begun to give more slots to disciplinary units that present candidate slates less skewed toward older white men. To reduce its geographic imbalance, NAS is also prohibiting members from nominating someone from their own institution. Today, 18 U.S. states have two or fewer members, whereas a handful of elite academic institutions each have more than 100.

BIOMEDICINE

Gene and cellular therapy trials take off

The number of new clinical trials for gene, cell, and RNA therapies has almost tripled in the past 4 years, the American Society of Gene & Cell Therapy said this month in its first ever quarterly trends report. As of 31 March, nearly 3500 of these experimental treatments were in development, most of which (53%) consist of altered genes or genetically modified (GM) cells, such as cancer-fighting T cells with modified receptors. Cancer is the target for the largest number (1200) of therapies under development. Among RNA therapies, many were vaccines (35) or COVID-19 treatments (30). The United States has more clinical trials underway in each of the three types of therapies than any other country: 1400 overall. Globally, 16 gene therapies (including GM cells), 53 non-GM cell therapies, and 15 RNA therapies have been approved for use so far.

Into the clinic

- Gene therapies
- Non–genetically modified cell therapies
- RNA therapies

(y-axis: 0, 100, 200, 300, 400, 500; x-axis: 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020)

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

A cancer patient in Louisville, Kentucky, receives a dose of a coronavirus vaccine.

COVID-19

Relief and worry for immune-suppressed people

Early studies suggest COVID-19 vaccine protection varies by ailment and treatment

By **Jennifer Couzin-Frankel**

For Eva Schrezenmeier, a nephrologist at Charité University Hospital in Berlin, the news was sobering: Among 40 patients with transplanted kidneys at her hospital who'd been vaccinated against COVID-19, only one was churning out the antibodies that would likely protect him from the disease. Because transplant patients take powerful drugs to suppress the immune system so it doesn't attack a donated organ, her team expected diminished responses to a vaccine. But Schrezenmeier, who posted a preprint describing her study last week, hadn't anticipated just how badly the vaccine might falter in her patients.

Her finding is at the grim extreme of research on how well COVID-19 vaccines work in the many millions of people whose immune systems are suppressed by drugs or disease. In many, the vaccines do seem to maintain their potency. But in others—particularly organ transplant recipients and those taking certain immune-dampening medications—effectiveness is less assured or even absent. To learn more, researchers are launching larger studies, seeking more clarity and ways to help patients whose weakened immune systems make protection against

COVID-19 all the more urgent. "There is a lot of confusion and fear among patients," says Alfred Kim, a rheumatologist at Washington University in St. Louis who cares for people with the autoimmune disease lupus and strongly urges vaccination for them.

One source of complexity: The dozens of different medications taken by people with cancer, autoimmune or other immunologic disease, or an organ transplant. Each can gum up different gears in the immune system's intricate machinery. The ailment makes a difference, too. Solid tumors such as colon cancer don't usually interfere with the immune system (although chemotherapy does). But autoimmune diseases or blood cancers such as leukemia and lymphoma can themselves deplete or disrupt certain types of immune cells.

Past research already suggested vaccines can falter in some immune-suppressed patients. Kim says flu and pneumococcal vaccines don't always work as well in people on some common immune suppressants, like methotrexate, which treats cancer and autoimmune diseases. And a 2012 study found that just 44% of cancer patients in treatment produced antibodies to influenza after one dose of flu vaccine; most were first vaccinated 1 week after chemo-

therapy. The researchers recommended two doses after finding that a second dose boosted the number to 73%.

When they started to parse blood samples after COVID-19 vaccination, scientists were unsure how people with immune suppression would respond to the vaccines. Gauging protection is also a challenge: The vaccines are designed to propel production of antibodies, but scientists don't know what levels are needed to guard against COVID-19. Antibodies are easier to measure than T cell responses, but those, too, play an important role in protection from disease.

Still, in a research setting, the hunt for antibodies can yield important clues. In December 2020, transplant surgeons Dorry Segev and Jacqueline Garonzik Wang at Johns Hopkins University put out a call on social media for organ recipients willing to participate in a COVID-19 vaccine study. "We had 1000 enrolled in the first week," Segev says. In March, the research team published details in *JAMA* of participants' immune responses to the first dose of the Pfizer-BioNTech and Moderna vaccines. The results foreshadowed Schrezenmeier's: Among 436 people who'd had liver, heart, kidney, and other organ transplants, just 17% had detectable antibodies.

Outcomes varied based on which medications the volunteers were taking, however. Only 9% of those on a class of drugs that includes the immunosuppressant mycophenolate had some antibodies, compared with about 40% in those not taking drugs in that category. Mycophenolate inhibits production of both B cells, which generate antibodies, and T cells, which help marshal B cells to do their job.

Segev says he and his colleagues are close to sharing results from his cohort's second vaccine dose, which show some improvement. Still, he's surprised that these organ transplant patients seem to respond even less well to COVID-19 vaccines than to flu vaccines. To learn more, he is studying their T cell, B cell, and other immune responses. "We're starting to try to say, 'What is going on here? Why is it so bad?'"

Although Segev worries about the roughly 500,000 transplant patients in the United States, he suspects the picture is much brighter for the 11 million people with autoimmune diseases, who tend to take different combinations of immune treatments or get by on lower doses. Last week, a paper in *Gastroenterology* reported that 48 people with either Crohn disease or ulcerative colitis, nearly all on immune-targeting medication, responded well to vaccination. Of the 26 whom the researchers followed through both vaccine doses, all produced antibodies, 22 at high levels.

But another study, of 133 people with various autoimmune diseases, suggested two types of medication can act as a sledgehammer against vaccine response. The work, posted as a preprint this month by Kim, rheumatologist Mary Nakamura at the University of California, San Francisco, and their colleagues, showed that on average, subjects churned out roughly one-third as many antibodies as healthy vaccinated people—a difference that doesn't strongly concern Kim. But people on therapies that destroy B cells, like rituximab, and the powerful steroid prednisone had far lower levels. Bigger studies of these patients are getting underway, including one announced last week by the National Institute of Allergy and Infectious Diseases.

In cancer patients, vaccine response likely depends at least partly on timing, because cycles of chemotherapy alternately squash immune cells and allow them to rebound, says Giuseppe Curigliano, an oncologist at the European Institute of Oncology in Milan. He reported last year that cancer patients on chemotherapy produced abundant antibodies after a bout of COVID-19, leaving him optimistic that vaccines will work well for them.

His center waits a couple of weeks after a chemotherapy cycle to offer a COVID-19 shot. Similarly, a U.K. study showed that, although many patients in treatment for solid tumors had a paltry response to the first vaccine dose compared with healthy volunteers, they appeared well-protected after the second. The researchers write that the results highlight the risks of delaying vaccine doses in cancer patients, contrary to the country's practice across its population.

There's nagging concern, though, when it comes to people with blood cancers. Ghady Haidar, a transplant infectious disease specialist at the University of Pittsburgh Medical Center, has preliminary results from patients with leukemia, lymphoma, and multiple myeloma suggesting a sizable fraction aren't producing antibodies after vaccination, particularly those with a form of chronic leukemia. Perhaps, he says, this occurs because patients "have defects in circulating white blood cells."

Physicians like Haidar say patients often ask whether to stop taking immune-suppressing medications before getting vaccinated, prompting tough choices. "No one should be stealth discontinuing meds so that they can respond to vaccines," he says. For some patients, skipping treatment can be dangerous, but doctors can sometimes delay an infusion of a therapy known to make a vaccine's job tougher.

For patients who don't appear protected by standard vaccinations, extra doses may help. Some organ recipients already get extra doses of hepatitis B vaccine, and this month, France recommended that they receive a third dose of the Pfizer-BioNTech COVID-19 vaccine. Christophe Legendre, a nephrologist at Necker Hospital in Paris, is planning antibody tests to see how well the approach works in transplant patients. Other researchers say labmade monoclonal antibodies might bolster protection for patients who still don't respond. (Although clinical trials have shown the monoclonal antibodies can prevent infection, so far they are only authorized for treating early-stage COVID-19.)

In Berlin, Schrezenmeier is planning to offer the AstraZeneca or Johnson & Johnson vaccines to some patients already vaccinated with another COVID-19 vaccine. Will mixing vaccines enhance their effectiveness? "I don't know," she admits. But she imagines that giving the immune system two different jolts might sometimes make a difference. The lone kidney transplant volunteer in her study who produced antibodies after vaccination had already survived COVID-19—which may have helped kickstart an immune response against it. ■

Science's COVID-19 reporting is supported by the Heising-Simons Foundation.

After revival, Iran's great salt lake faces peril

Upcoming presidential election could stall efforts to restore Lake Urmia

By **Richard Stone**

Twenty years ago, geochemist Arash Sharifi began to drill sediment cores in Iran's Lake Urmia—then the largest lake in the Middle East—to probe its recent climate history. "I was shocked at how little was known about the lake," recalls Sharifi, now at Beta Analytic Inc. in Miami. He became entranced by the "very unique chemistry" of its hypersaline waters. He also grew alarmed: Dams on feeder rivers and a proliferation of illegal wells had made the lake, a favorite haunt of flamingos and migratory birds, "vulnerable to hydrological collapse," he wrote in an internal government report.

That dark vision came to pass: By 2013, Lake Urmia's surface area, once 6100 square kilometers, or about the size of Delaware, had shrunk by 90%, a plight reminiscent of that of the larger Aral Sea in Central Asia. To save the dying lake, Iranian President Hassan Rouhani launched a restoration initiative that has so far cost $1 billion. Now, Lake Urmia has a stronger pulse: Water levels have risen, and last summer flamingos were spotted in large numbers for the first time in years.

But the rebound could be fleeting. The higher water appears to be almost entirely the result of a few years of unusually strong rains, a recent study concludes, but that pattern is already giving way to drier conditions. And Iran's June presidential election is expected to usher in a leader less committed to the lake's restoration than Rouhani, who can't run for reelection under the nation's constitution. "We're at a tipping point," says Somayeh Sima, a water resources specialist at Tarbiat Modares University in Tehran, Iran. Even pausing restoration work for a year or two, she says, "would be dangerous for the lake."

The shallow lake was at its most expansive in recent decades in the late 1990s, when it stretched nearly 140 kilometers

Iran's government hopes to save Lake Urmia for both wildlife and recreation, including therapeutic baths in its very salty waters.

long and 50 kilometers wide, and its volume exceeded 30 billion cubic meters. But a 5-year dry spell that began in 1998 was coupled with increasing water diversion for irrigation. "We call it anthropogenic drought," says Amir AghaKouchak, a hydrologist at the University of California (UC), Irvine, who first visited Urmia as a student in 1998. By 2013, Urmia's volume had fallen by a factor of 60, to 500 million cubic meters, exposing salt flats that fueled salt-ridden dust storms.

Under Rouhani's government, which took office in 2013, providing water for drinking, irrigation, and industrial uses has remained paramount. But in 2018 the Cabinet mandated that managers start to allocate some water to wetlands. The government also launched the Urmia Lake Restoration Program (ULRP), which has endeavored, with mixed success, to wean nearby farmers off thirstier crops like sugar beets and adopt more efficient irrigation practices. ULRP managers have also played whack-a-mole with illegal wells, pouring concrete into some as others pop up.

More controversially, ULRP has emphasized big engineering works, including a dam, tunnel, and canals that divert as much as 600 million cubic meters of water a year from the Little Zab River to Urmia's south arm. (Like Utah's Great Salt Lake, Urmia is bisected by a causeway.) Together with treated wastewater funneled into the lake, the inputs may have helped inundate lakebed that had dried out several years ago, maintaining the isolation of two islands that are home to the rare Persian fallow deer and Armenian sheep. (At the lake's nadir, predators could walk across dry lakebed from the mainland to the islands.) However, part of the Little Zab watershed is in Iraq, in territory controlled by Kurdish communities with their own water needs. As upstream communities tap the water, AghaKouchak believes, the flow into the lake is "bound to suffer in the future."

In the past few years, such worries receded as ample precipitation helped boost lake levels by about 1.5 meters. "Nature has been kind," says Soroosh Sorooshian, a water expert at UC Irvine. The rains benefited water bodies across the broader region, including Lake Van in Turkey and a pair of reservoirs in Iraq, a team led by Behzad Ataie-Ashtiani, a civil engineer at Sharif University of Technology, reported this month in the *Journal of Great Lakes Research*. But historically, droughts have followed rainy periods in the region, and this year has so far been much drier. Those vagaries make it hard to judge whether the restoration program is succeeding.

It's also not clear that ULRP can undo damage already done by the lake's shrinkage. The project aims to raise Urmia's levels by an additional 3 meters, which planners say would reduce the lake's salinity from its present supersaturated state—about 350 grams of salt per liter of water—to between 240 and 285 grams per liter. At that level, brine shrimp—now present in vanishingly low numbers in the lake—could return at great enough densities, ULRP managers hope, to nourish flamingos and other wildlife. (The recent flamingo sightings suggest the birds can subsist on other critters as well.)

But some experts doubt the target lake level will be sufficient to dilute meters-thick salt deposits, accumulated when now-submerged swaths of lakebed dried out several years ago. Still, Sima notes, any increase in lake levels is likely to benefit ecosystems and people.

Such technical debates could become moot depending on the outcome of the June election. With Iran's economy crippled by Western sanctions, the country's next president, widely expected to be a hardliner, will be hard pressed to steer scant resources to environmental protection. "Nobody wants to see the lake die," Ataie-Ashtiani says. "But the situation is very fragile." ∎

Richard Stone is senior science editor at the Howard Hughes Medical Institute's Tangled Bank Studios in Chevy Chase, Maryland.

IRAN
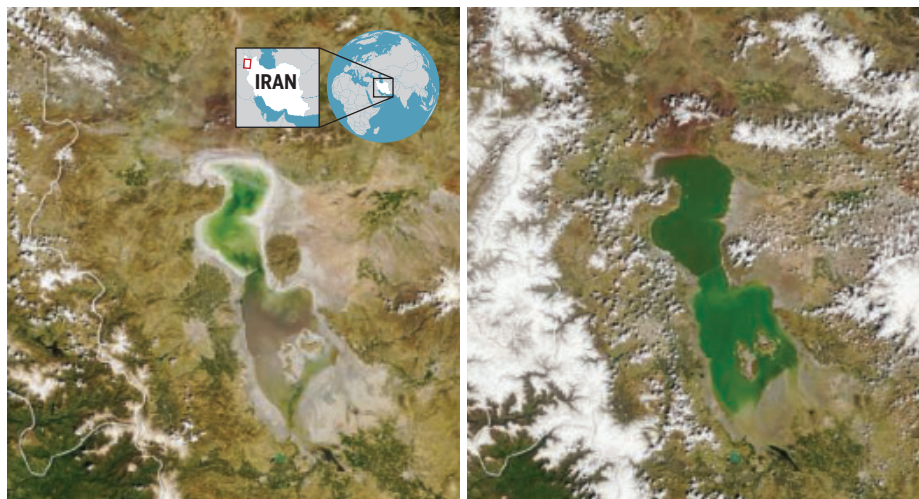
Lake Urmia had shrunk calamitously by 2014 (left), then roared back in 2019—but its health is still precarious.

ENERGY TECHNOLOGY

Gravity powers batteries for renewable energy

With winches, weights, and disused mine shafts, company looks to beat chemical batteries

*By **Cathleen O'Grady**, in Edinburgh, U.K.*

Alongside the chilly, steel-gray water of the docks here stands what looks like a naked, four-story elevator shaft—except in place of the elevator is a green, 50-ton iron weight, suspended by cables. Little by little, electric motors hoist the weight halfway up; it is now a giant, gravity-powered battery, storing potential energy that can be released when needed. And that moment is now: With a metallic moan, the weight inches back down. Reversing direction, the motors become electric generators, sending up to 250 kilowatts of power back to the grid. For peak power, the weight can descend in 11 seconds—but for testing purposes, it moves just a few meters at "creep speed," says Douglas Hitchcock, project engineer at Scottish startup Gravitricity.

The company announced last week that its small-scale demonstrator is now operational, capable of switching between drawing energy from the grid and sending it back in seconds. The design offers an alternative to the chemical batteries that dominate the global energy storage market—a market that is growing hand in hand with renewable power, which needs to bank energy when the Sun shines or the wind blows, and release it when the grid faces high demand.

Gravitricity is one of a handful of gravity-based energy storage companies attempting to improve on an old idea: pumped hydroelectric power storage. Engineers would dam up a reservoir on a hill, pump water to it at times of low demand (usually at night), and release it to generate electricity. But the systems require specific terrain, expensive infrastructure, and planning approval that is increasingly hard to come by. These days, banking energy usually means hooking up renewable power to giant batteries.

Yet gravity-based storage has some distinct advantages, says Oliver Schmidt, a clean energy consultant and visiting researcher at Imperial College London. Lithium-ion batteries, the technology of choice for utility-scale energy storage, can only charge and discharge so many times before losing capacity—usually within a few years. But winches, steel cables, and heavy weights can hold up well for

decades. "It's mechanical engineering stuff," Schmidt says. "It's relatively cheap." And mining the minerals for batteries brings environmental and human rights problems, says Miles Franklin, Gravitricity's lead engineer; a bucket of iron has a much lighter footprint.

Using Gravitricity's own cost and performance estimates, Schmidt compiled a 2019 report for the company showing that all told—including construction, running costs, and maintenance—gravity storage can be cheaper than lithium-ion batteries. For a 25-year project, he estimates Gravitricity would cost $171 for each megawatt-hour. Jessika Trancik, an energy storage researcher at the Massachusetts Institute of

A towering, low-tech "battery" stands on the docks in Edinburgh, U.K.

Technology, says that number still needs to be supported with field data. But Schmidt's calculation of the lifetime cost per megawatt-hour for lithium-ion batteries, $367, is more than twice as much, although battery costs are dropping. Flow batteries, a promising grid-scale technology that stores charge in large tanks of liquid electrolyte, come in at $274 per megawatt-hour.

Other gravity-based storage companies have their own twists on the technology. The idea behind California-based Gravity Power is just a small step away from pumped hydro: It uses renewable energy to pump water under a heavy piston and lift it. When power is needed, the piston weight is released, forcing the water through a hydroelectric generator. The German company New Energy Let's Go uses a similar design. Switzerland-based Energy Vault wants

to use a multiarmed crane with motors-cum-generators to stack and disassemble a 120-meter-tall tower made of hundreds of 35-ton bricks, like a Tower of Babel that rises and falls with energy demand. Gravitricity keeps it simple, with its elevator-like shafts. It plans to scale up to 500-ton weights, which would require mammoth foundations for a tower—so the best place to put a full-scale system is underground, says business development manager Ruth Apps. The company is scoping out disused mines in the Czech Republic, Poland, and South Africa for its first commercial projects.

The technology is still "incredibly immature," Schmidt cautions, and companies have made slow progress. Energy Vault, probably the leader, announced in 2019 that it had raised $110 million and plans to start commercial developments this year. But like all storage technologies, gravity-based storage will flounder if climate regulations don't create incentives for carbon-free energy, says Rebecca Willis, an environment policy researcher at the University of Lancaster. In most places, she says, small natural gas plants that can be easily turned on and off remain the cheapest way to cope with fluctuations in demand.

With a staff of 14 and just £3 million in investment, Gravitricity has no illusions about the hurdles ahead. As the team put the demonstrator through its paces, they have already turned up unexpected difficulties—like the torques exerted by the steel cables, which want to untwist as they lift the weight. After working through these teething issues, the company by 2023 plans to build a full-scale plant—with the heavier weights and a shaft nearly a kilometer deep that could produce up to 4 megawatts of peak power.

Ghosts of an ailing industry surround Gravitricity's test site: A supply vessel serving North Sea oil rigs slumbers nearby. But on a clear day, one can make out Scotland's energy future: the wind turbines that dot the horizon less than 20 kilometers from here. Two of Gravitricity's recent hires come from the oil and gas sector, Apps says. And mining communities are thrilled at the prospect of turning their heritage to something new: "They see the idea of having a second life." ∎

As part of Fieldlab, 5000 fans attended a 27 March soccer match between the Netherlands and Latvia.

COVID-19

Dutch studies bring back the fun—but are they good science?

Critics question protocols and ethics of experimental gatherings that rely on testing to keep people safe

By **Jop de Vrieze**

The Eurovision Song Contest, known best for its over-the-top performances and outrageous costumes, has a new feature this year: It will be the site of a massive field experiment to see whether concerts and other events can be held safely in the middle of the COVID-19 pandemic. Nine rehearsals and televised shows, staged 18–22 May in Rotterdam, Netherlands, will each be attended by 3500 visitors who will have to show a recent negative SARS-CoV-2 test to get in. Those admitted can choose to drop social distancing and go without face masks—precautions currently mandatory in indoor public spaces in the Netherlands, where most people remain unvaccinated.

The contest will be the last of 20 experimental events, together dubbed Fieldlab, set up by the Dutch event industry in collaboration with scientists and the Dutch government. But Fieldlab has come under fire as events have grown bigger and COVID-19 cases in the Netherlands surged. A music festival for 10,000 people on 24 April was banned by the host city, Breda, after more than 300,000 people signed a petition opposing it. And last week, more than 350 researchers criticized the studies in a letter that complained of a lack of peer review,

an intransparent setup, and ethical failings.

"Basic conditions and standards for scientific research do not appear to have been met," the authors wrote. "A festival with 10,000 visitors … is not risk free, even with entrance testing," says Caspar van Lissa, a methodologist at Utrecht University, who wrote the open letter. "If it were, there would be no need to do the study."

Fieldlab's goal is "to determine what's an acceptable risk for visitors, event organizers, and administrators," says Bas Kolen, a security researcher at the Delft University of Technology involved in the study. The first two events—a theater show and a business conference, each with 500 attendees—took place in February. The researchers found that, with the virus prevalence at that time, pre-event testing and additional measures such as ventilation could keep attendees' risk at about one infection per 100,000 people per hour—the same risk they would run by staying home. Bigger events followed, including a soccer match between the Dutch and Latvian national teams with 5000 fans.

The studies didn't need approval from a medical ethics committee because they didn't meet the legal definition of medical research, a panel at Radboud University Medical Center ruled. But the authors of the open letter say Fieldlab should have followed ethical guidelines for research in the social and be-

havioral sciences, which stipulate that participants give their informed consent and researchers assess the potential drawbacks for individuals and society. "Not a single behavioral scientist is involved. If they were, this would have never happened," says psychologist Denny Borsboom of the University of Amsterdam. Andreas Voss, an infectious disease specialist at Radboud University who leads the project, notes that the social science guidelines are not mandatory and that tickets came with conditions saying Fieldlab could not be held liable for infections.

The critics also question Fieldlab's statement that the events are, on the whole, safe. Participants are requested to take a second COVID-19 test 5 days after the event, and at least 25 people have tested positive, although for most of them it's hard to determine whether they were infected at the event. Fieldlab's main gauge of risk isn't the number of infections detected, however, but the number predicted by a model that incorporates data on ventilation and people's behavior, including mask wearing and how many superficial or close, extended contacts they have at the event, measured using trackers and video analyses.

Kolen, whose team specializes in calculating the risk of floods, concedes the model has many assumptions and limitations. It assumes testing before each event will identify 95% of infectious individuals, for example, and it does not take into account the possibility that an infectious person happens to be a social butterfly, or sheds very large amounts of virus. "Those are characteristics we would like to explore in a later stage," Kolen says. "It's what everyone does who is modeling infectious diseases," Voss adds. But to Van Lissa, "These are exactly the factors that you should empirically investigate to know how safe an event is." In response to the criticism, the Fieldlab team has published most of its protocols and a description of the risk model. Kolen says he supports "a healthy debate about its value."

Some scientists see value in the project. Although "the scale and the timing" of the events might be "unfortunate," virologist Marion Koopmans of Erasmus Medical Center tweeted last week, "the basic design is relevant, and the data collected will provide input for modelers for years."

Whether Eurovision will welcome visitors is now unclear: Its organizers say they are closely following public sentiment and epidemiological trends before making a definitive decision. ∎

Jop de Vrieze is a science journalist in Amsterdam.

INFECTIOUS DISEASE

Malaria vaccine achieves striking early success

But observers want to see results from larger, phase 3 African trial

By **Meredith Wadman**

After decades of disappointing results, new findings have revived hopes for an effective vaccine against malaria, which kills some 400,000 people every year, most of them children. An experimental vaccine that targets the most dangerous form of the malaria parasite was found to have an efficacy of 71% to 77% after 1 year in children.

The results, posted as a preprint last week, come from a trial of a vaccine developed by researchers at the University of Oxford's Jenner Institute involving 450 toddlers in Burkina Faso, where malaria is endemic. "The efficacy we have got has never been obtained by any

Children in Burkina Faso receive medicine aimed at preventing malaria.

[malaria] vaccine candidate. These are really amazing findings," says trial site co–principal investigator Halidou Tinto, a parasitologist at the Institute for Health Sciences Research in Nanoro, Burkina Faso.

The results are "very positive news," says Pedro Alonso, director of the World Health Organization (WHO) Global Malaria Programme, who was not involved with the work. But he and others note the trial's small size and that the vaccine's protection was really only demonstrated during the 6 months when malaria was most prevalent in Burkina Faso; scarcely any malaria cases occurred at other times. "We are still quite far away from having the type of information that would allow us to get very excited," he says.

But the new study's investigators are bullish and plan to launch a pivotal phase 3 trial later this year, enrolling 4800 children in Burkina Faso, Mali, Kenya, and Tanzania. In the best case, data from that trial could be submitted to regulators late in 2022 for approval in early 2023, says Oxford's Adrian Hill, the trial's chief investigator. He hopes that the phase 3 results will persuade regulators to issue emergency use authorizations, as they have for COVID-19 vaccines.

The children in the current trial, aged 5 months to 17 months, received three doses of vaccine at 4-week intervals and a booster at 12 months. Of the 146 children whose vaccine included a high dose of an immune-boosting compound called an adjuvant, 39 developed malaria, versus 106 of 147 children in a control group who received rabies vaccine. (The rabies shots ensured the control group also received value from being in the trial.)

The 77% efficacy against malaria dipped to 71% in children who got a vaccine with a lower dose of adjuvant. The children's levels of specific antibodies to malaria fell by two-thirds by 9 months, but the booster dose at 12 months restored them, according to the paper, which is in press at *The Lancet* and was posted 20 April on its preprint server.

WHO has called for the development of vaccines that can reduce malaria cases by 75% by 2030. But the malaria parasite's complex life cycle and shifting surface proteins have challenged vaccine developers. The

highest efficacy previously published for a vaccine at 1 year after dosing was 56%, for Mosquirix, made by GlaxoSmithKline (GSK), Hill notes. Mosquirix's efficacy in a large trial dropped to 36% after about 4 years, and some scientists worry the same could happen with the Oxford vaccine: The vaccines are structurally similar and both target the parasite right after infection.

Hill believes his team's vaccine will have staying power. The Oxford vaccine consists of hepatitis B protein combined with part of a protein that coats the surface of the malaria parasite. The hepatitis B proteins self-assemble into a "viruslike particle" densely studded with malaria proteins—more densely than in GSK's vaccine. Hill thinks this amplifies antibody responses and protection.

The Serum Institute of India is making the Oxford vaccine for a planned phase 3 trial and has pledged to produce 200 million doses in the coming years. "The price [will be] less per dose than ever before," Hill says. The potent adjuvant, a purified plant compound called a saponin, made by Novavax, is already being manufactured at scale for that company's COVID-19 vaccine.

Others caution that many unknowns remain. In many countries, malaria transmission is continuous, not seasonal, notes Robert Sauerwein, a medical microbiologist at TropIQ Health Sciences in Nijmegen, Netherlands. "The real issue, sustained protection and induction of sustained [immunological] memory, has not been addressed yet," he says. (The investigators are currently following the children in the trial out to 24 months.)

Others called out what they saw as deficiencies in the study. "Where's the biology?" asked Rhoel Dinglasan, who is developing another malaria vaccine at the University of Florida's Emerging Pathogens Institute. He applauded the new results, but wanted to see genomic data on the malaria parasites that managed to infect vaccinated children.

Dinglasan says that without sequence data, it's not clear whether the vaccine will be effective against all parasite strains. The efficacy of the GSK malaria vaccine fell by one-third, to 33.4%, when there was a mismatch between a key part of the protein used in the vaccine and the corresponding protein in the parasite, he notes. And a mismatched vaccine could drive selection for vaccine-resistant parasites, potentially thwarting the quest for a highly effective vaccine once again. ∎

A forest gives way to a plantation on Sumatra. Deforestation and other major challenges mean that Indonesia needs science more than ever, researchers say.

RESEARCH POLICY

'Superagency' may further politicize Indonesian research

New National Research and Innovation Agency will have broad powers to fund, execute, and control science

By **Dyna Rochmyaningsih**

Indonesia has dismantled its science ministry and created an overarching national research agency, a move some scientists worry will strengthen political control over research in a country where academic freedom is already under pressure and politics have taken an authoritarian turn.

The Indonesian Parliament on 9 April approved a proposal by President Joko Widodo to eliminate the Ministry of Research and Technology (RISTEK) and create a new National Research and Innovation Agency (BRIN). Details have yet to be fleshed out, but BRIN seems set to have broad powers to fund, execute, and control research in the country. It will be led by physicist Laksana Tri Handoko, who currently heads the Indonesian Institute of Sciences (LIPI).

Widodo has often criticized the Indonesian scientific community for what he says is lackluster performance. The country's $1.7 billion annual research budget—a fraction of what the United States and many European countries spend—is enough, he told Indonesian academics in 2019. "Where's the output?" he asked. Widodo has also been critical of the large number

of research agencies scattered around the national bureaucracy and provincial governments across the archipelago.

BRIN, which Widodo has called "the big house" for Indonesian research, was formally created in a new science law adopted in 2019, but debates about its role and place in the Indonesian government delayed it. Some initial plans called for it be part of RISTEK, but Widodo decided BRIN will replace it and several other scientific bodies, including LIPI and the Agency for the Assessment and Application of Technology. ("Personally, I'm sad because I am the last research minister in the country," RISTEK Minister Bambang Brodjonegoro said at an 11 April seminar.) Besides planning and funding research, the new "superagency" will also issue permits for foreign researchers to work in Indonesia—an increasingly contentious topic.

BRIN was conceived within Widodo's political party, the Indonesian Democratic Party of Struggle (PDID), whose leader, former President Megawati Sukarnoputri, has warned against foreign interference in research. "Don't let our republic rush into collaborating with our close friends from abroad," she told *The Jakarta Post* in January 2020. "Just make use of ... our own people first." Some see Sukarnoputri's

ambitions as an extension of the policies of her father, Sukarno, who as Indonesia's first president in the 1950s and '60s promoted science and technology as a way for the newly independent nation to develop.

Indonesian politicians, from PDID and other parties, "want our science and technology to be in line with their nationalistic ideology," says Berry Juliandi, a spokesperson of the Indonesian Young Scientists Forum. Although he says there is nothing wrong with striving for national greatness, "We have to make sure that there will be no politicization."

Yet political influence on research is growing, says Inaya Rakhmani, a media sociologist at the University of Indonesia. The government has clamped down on unwelcome research into deforestation, forest fires, and threats to a rare orangutan population, for example. Historians are not allowed to study anticommunist massacres in 1965 and 1966 during which up to 1 million Indonesians perished, and human rights scholars are banned from advocating for independence for West Papua.

Yanuar Nugroho, an adviser for the Center for Innovation Policy and Governance, says the reforms in Indonesian research policy come at a time when the country needs solid science more than ever. A plan to create massive new crop plantations on peatlands and in protected forests on Borneo will further increase the risk of devastating wildfires and floods, scientists say. The government is also supporting an unorthodox vaccination method for COVID-19 that involves taking dendritic cells from a person's body, immersing them in viral antigens, then reinjecting them. The Indonesian Food and Drug Authority, worried about side effects and unsterile production procedures, has banned a phase 2 clinical trial of the method, developed by former Minister of Health Terawan Putranto, who also invented an unproven and potentially harmful stroke therapy.

Nugroho hopes BRIN's power can still be limited by making it responsible only for coordinating and carrying out research, while making another agency responsible for policy. Juliandi would like to see funding put into the hands of an independent organization as well. "When you centralize everything in one governing body, there is always a chance for abuse of power," he says. ∎

Dyna Rochmyaningsih is a science journalist based in Deli Serdang, Indonesia.

THE SENTINELS

In Brazil, a team of scientists watches the world's largest rainforest for pathogens that could spill from animals to people

*By **Dan Grossman**; Photography and reporting from Manaus, Brazil, by **Dado Galdieri** of Hilaea Media*

In a lab set up in a Manaus, Brazil, rainforest park, Aline Ramos (center) and colleagues collect samples for a biobank from a pied tamarin monkey.

Pied tamarins living within the city of Manaus, Brazil, could be a disease risk for humans, but human infections such as Zika may threaten the monkeys, which are in decline.

When Marcelo Gordo opens the picnic cooler, the stench is suffocating. Three dead pied tamarin monkeys, their cream-and-caramel-colored coats visible through plastic wrap, are curled up inside. Gordo, a biologist at the Federal University of Amazonas, Manaus, explains that a student accidentally unplugged the freezer where he'd stored the monkeys, which had been killed on the road and given to him by city officials. Despite the decay, they are worth investigating.

Inside the spartan necropsy room at a veterinary school here, veterinarian Alessandra Nava and two graduate students pull on goggles, N95 masks, and blue nitrile gloves and begin to cut bits of tissue and collect bodily fluids from the monkeys. They pack the samples into vials to be transported to the Fiocruz Amazônia Biobank, a pathogen research collection that Nava helps oversee at the Amazonian regional office of the Oswaldo Cruz Foundation, a branch of Brazil's Ministry of Health more commonly known as Fiocruz. There, she and others will test the samples for parasitic worms, viruses, and other infectious agents.

Nava and her colleagues are on the front lines of the search for animal diseases that

could spill over and infect humans—and perhaps cause the next pandemic. New diseases can come from anywhere: Severe acute respiratory syndrome and COVID-19 both originated in China, for instance. Another recent coronavirus disease, Middle East respiratory syndrome, was first found in Saudi Arabia. But many researchers suspect tropical rainforests, with their stag-

Roundworms infested a monkey that was examined at a wildlife rescue center in Manaus.

gering biodiversity, are the most likely cradle of dangerous new pathogens.

When human populations encroach on rainforests, the risk of spillover skyrockets. Manaus, Brazil, a city of 2.2 million people in the Amazon rainforest, is just such a place. The jungle that stretches for hundreds of kilometers in every direction has long threatened inhabitants with infections circulating in wildlife. Some 12% of the word's 1400 bat species—known to host a bewildering range of viruses—flit through the Amazon forest. Its monkeys and rodents carry plenty of potential threats as well.

Urban growth, highway expansion, hydroelectric dam construction, mining for gold, and deforestation for cattle ranches and small farms erode the jungle and bring humans and wildlife into ever closer contact. In Brazil, the pro-business policies of President Jair Bolsonaro have only boosted that risk. By monitoring local animal populations and human patients, researchers at Fiocruz hope to head off zoonoses—diseases that leap from animals to humans—before they spiral out of control. Their work highlights the importance of curbing human activities that boost the risk of spillover. It could also guide surveillance for new and rare diseases in hospitals, which would enable health workers to respond fast if a rainforest pathogen became a wider threat.

Ironically, Fiocruz's work has been stymied by one such disease. Manaus has experienced two brutal waves of COVID-19, a disease thought to have originated in bats. The city's cumulative death toll, roughly 9000, is among the world's highest per capita. Nava's team has not captured animals at field sites in a year, partly out of concern that the researchers themselves might infect wild animals with the coronavirus. And the labs at Fiocruz Amazônia that process her samples have been commandeered for coronavirus research.

For Felipe Naveca, the lab's vice director of research and innovation, the upheaval has been personal as well as professional. COVID-19 killed his father and may have contributed to the death of his grandmother. In the lab, Naveca led one of the first genetic studies of the new P.1 coronavirus variant that has emerged from Manaus and appears to be especially dangerous because it is more transmissible and evades immunity. He is proud that his team has processed 18,000 COVID-19 tests for local health authorities. "Helping to save someone's life was much more rewarding than publishing a scientific article," he says. But like his colleagues, Naveca is anxious to get back to the lab's core mission. "We must keep searching for those emergent threats."

THE OFFICE of Fiocruz Amazônia occupies a former military hotel in downtown Manaus, nestled between a small church and a luxury condominium high-rise. Several rooms with softly humming freezers and refrigerators house the biobank: a collection of feces, blood, and other tissues and fluids from more than 100 rainforest animals. Forty species are represented; the majority are monkeys, bats, and rodents, the mammals thought most likely to transmit disease to people. Other collections in the building include insects that torment these animals and could serve as vectors for ferrying pathogens to humans.

The Fiocruz Amazônia Biobank was partly modeled on the $200 million PREDICT early warning program. Launched in 2009 by the U.S. Agency for International Development, PREDICT identified nearly 1000 previously unknown animal viruses with zoonotic potential before the Trump administration canceled it in 2020. Whereas PREDICT was global, Nava and her colleagues do the same kind of work at a regional level. They're searching for animal reservoirs of known viral and parasitic diseases, including obscure viral fevers and filariasis, a parasitic worm infection that can cause the horribly disfiguring syndrome elephantiasis. They're also using DNA sequencers to scour samples from animals for pathogens that have yet to emerge.

"What they are doing is brilliant and important," says Andrew Dobson, a biologist at Princeton University who studies the ecology of wildlife diseases. "It shows that even in countries with limited resources and a very negative governmental attitude towards science, it is possible to set up a monitoring scheme for novel viruses."

Veteran disease hunter Dennis Carroll, who founded and ran PREDICT, agrees. "Amazonia is one of the richest, most ecologically diverse regions of the world," he adds. "So getting any insight into that region is really important."

Students wait to sample animals while veterinarian Alessandra Nava stores specimens in liquid nitrogen (right).

A black-and-white photograph in a second-floor foyer of Fiocruz Amazônia depicts one inspiration for this work: Brazil's legendary doctor and disease sleuth, Carlos Chagas. Attired for an expedition in a white suit and knee-high boots, Chagas stands in a canoe surrounded by his oarsmen. In 1909, Chagas discovered the cause of the disease that now bears his name. Using a simple microscope, he identified the culprit as a protozoan (now called *Trypanosoma cruzi*) and showed that it is transmitted by the bite of triatomine bugs, often called kissing bugs. Chagas disease, whose symptoms range from fever to heart failure decades later, still kills hundreds of thousands of people a year in Latin America.

Naveca is doing similar detective work with the more sophisticated tools of modern genetics. One pathogen that concerns him is the little-studied Oropouche virus, which is spread primarily by a species of midge, *Culicoides paraensis*. Oropouche, which causes fever, headache, and joint pain, has sparked at least 30 outbreaks and sickened more than 500,000 people since it was first identified in 1955. Its range has gradually expanded to include Panama, six South American countries, and Trinidad and Tobago, where it first appeared. The midge itself, however, lives as far away as the northern United States, where it and related insects are called no-see-ums, sug- gesting the virus could spread beyond South America. The southern house mosquito (*Culex quinquefasciatus*), a carrier of West Nile and Saint Louis encephalitis viruses, can also transmit Oropouche, though not very efficiently, and its range throughout the tropics raises the possibility of Oropouche outbreaks in Africa, southeast Asia, and Australia.

Naveca and his colleagues hope to find out which animal or animals are the primary natural reservoirs for this virus. There are plenty of candidates: Oropouche has been identified in sloths, marmosets, finches, and several other birds and mammals. The team recently reported using the polymerase chain reaction to identify

Illegal construction encroaches on rainforest near Manaus, turning people and animals into intimate neighbors.

the virus' genetic material in urine and saliva—as opposed to blood—which could make the hunt for its animal reservoir easier and aid diagnosis in patients.

Naveca is also worried about another little-studied virus that is rapidly expanding in South America: the Mayaro virus, which causes flulike symptoms, making it hard to distinguish from more common tropical diseases such as chikungunya and dengue fever. As with Oropouche, he's hoping to pinpoint the virus' natural reservoirs and investigate whether cases of it are going undiagnosed.

Mayaro is a likely candidate for the next large-scale outbreak of an animal virus in Brazil or beyond, Naveca and other scientists warn. Its primary vector, the mosquito *Haemagogus janthinomys*, is a forest dweller restricted to Central America and northern South America, but laboratory experiments show the yellow fever mosquito (*Aedes aegypti*) and the Asian tiger mosquito (*A. albopictus*)—two species widely distributed in tropical and subtropical

areas—can also transmit the disease. *A. aegypti* is especially well adapted to breeding in cities.

To Naveca, the Zika virus is a case study in the value of tracking obscure pathogens. First identified in Africa in 1947, where it spilled over from monkeys, it circulated largely unnoticed and with few casualties for decades. Then, it caused an outbreak in Oceania in 2013 and, 18 months later, a massive epidemic in Latin America. Researchers suddenly discovered a disturbing consequence of the disease—microcephaly and other birth defects in infants born to infected mothers. "Zika was a virus that nobody was paying attention to until 10 years ago," Naveca says. "We can fight better the enemies we know better."

Naveca now hopes to carry on Chagas's disease-hunting tradition with a deal he's negotiating to procure a 25-meter, flat-bottomed boat that has been outfitted to be a floating laboratory. Preserving perishable human and animal samples at remote field

sites has been a critical obstacle, and the vessel would bring the lab to the biological materials, rather than the other way around. Naveca hopes to join its maiden research voyage, possibly later this year, to remote Amazon villages, where he and colleagues plan to trap bats, rodents, primates, and insects, and bring a trove of specimens back to Fiocruz Amazônia.

EVEN WITHIN MANAUS, there are lots of opportunities for fieldwork. When *Science* visited last year, Gordo had set up an improvised lab inside a classroom at Sumaúma State Park, a tiny patch of uncut rainforest in the middle of the city, wedged between a busy highway and an upscale mall. Using cages baited with ripe bananas, he and his assistants trapped nine pied tamarins and injected them with a sedative, then swabbed their oral and anal cavities, clipped locks of hair, and drew blood. Then they set the animals free.

It's peculiar and sometimes dangerous

Urban jungle

Manaus, Brazil, with more than 2 million people, lies at the heart of the world's largest rainforest, making it a good spot for tracking the forest's pathogens—as well as a potential spillover point.

work. Monkeys have bitten and sneezed on Gordo, and on this trip a syringe broke as he squeezed the plunger, spraying monkey blood on his face shield. He says his wife complains when he stashes monkey carcasses in their home fridge.

Manaus's Yoda-faced pied tamarins live all over the city. Like North American squirrels and raccoons, they don't respect property lines and make urban gardens their pantries and playgrounds. There's no evidence so far that Manaus's urban monkeys are a human health threat, and Gordo, worried about "unreasonable killings or deforestation," is reluctant to discuss that possibility. But he and others are investigating whether monkeys carry parasites, such as the nematodes that cause filariasis, or viruses such as Zika and chikungunya.

For Gordo, an equal concern is spillback—infections passed from humans to wildlife. Zika, for example, appears to have traveled from humans back to wild monkeys during Brazil's epidemic. Fears that the virus might harm wildlife rose when researchers showed that a pregnant monkey native to Brazil had a spontaneous abortion after it was exposed to Zika. The fetus had birth defects similar to those seen in humans.

So far, Gordo has not found the virus in Manaus's monkeys, but they may be at risk: A study he co-authored last year found mosquitoes from two species thought to carry Zika, *Haemagogus janthinomys* and *Sabethes chloropterus*, in both monkey and human habitats in a forest reserve on the edge of the city. The pied tamarins are already critically endangered, found nowhere else but in and around Manaus. Their population is expected to decline by 80% within the next 16 years. A virus outbreak could push them over the edge.

Humans are at risk from spillback as well. In Europe and the United States, scientists worry about COVID-19 outbreaks on mink farms, for example, because such events give the virus more opportunities to evolve and jump back into people. Likewise, primate populations infected with Zika could reignite human outbreaks. This happened with yellow fever: Brought to South America centuries ago with the slave trade, the virus has been impossible to eliminate from Brazil because it established itself in wild monkey populations, which occasionally pass it back to people.

After trapping monkeys for a day in the Sumaúma park, Gordo went home and bottle-fed an infant pale-throated sloth only slightly larger than his cupped hands. A friend had found it untended on the ground in a forest fragment not far from his university office. Despite everything he's learned about zoonotic diseases, Gordo said he was "not too worried." The sloth pup looked healthy. But several weeks later it got sick and died, possibly from pneumonia.

NAVA BELIEVES the Fiocruz center's work is only becoming more urgent with changing land use patterns in the Amazon. Deforestation has soared since Bolsonaro came to power in 2019—transforming habitat in ways that could make viral hosts and vectors more dangerous and increasing the likelihood of spillover.

In 2016, she and colleagues reported that 9% of bats in small clearings around settlements in Brazil's coastal Atlantic Forest had active infections of one or more of 16 viruses, including coronaviruses and hantavirus. In less-disturbed forests nearby, fewer than half as many bats were infected, and with only six different viruses. The findings fit a widely debated hypothesis known as the dilution effect, which holds that in forests with greater biodiversity, mosquitoes and other vectors have more targets and end up biting animals not capable of incubating a given virus, thereby slowing its spread. Reducing biodiversity by clearing land can do the opposite, and it also pushes humans into closer proximity to wildlife. Bats are a particular concern, Nava says, because they often roost in buildings.

It all underscores the need to stop destroying rainforest, she says—although she acknowledges that Brazil's policies are unlikely to change under Bolsonaro, who has nearly 2 years left in his term. In the meantime, Nava says, disease fighters must keep monitoring the jungle for dangerous diseases. "We have no power to reduce deforestation," she says. But, she adds, "We have the power to search for new viruses." ∎

Dan Grossman is a journalist in Watertown, Massachusetts. Photojournalist Dado Galdieri is based in Brazil.

INSIGHTS

Florida torreya (*Torreya taxifolia*) is a species native to Florida that is currently undergoing unregulated assisted colonization to other US states.

POLICY FORUM

CONSERVATION

Global policy for assisted colonization of species

Coordinated policies are needed for the translocation of species for conservation

By **Jedediah F. Brodie**[1,2], **Susan Lieberman**[3], **Axel Moehrenschlager**[4,5,6], **Kent H. Redford**[7,8,9], **Jon Paul Rodríguez**[10], **Mark Schwartz**[11], **Philip J. Seddon**[12], **James E. M. Watson**[13,14,15]

N egotiations in advance of the 15th meeting of the Conference of the Parties to the Convention on Biological Diversity (CBD) (*1*) in October 2021 will set the course of international conservation for the next several decades, providing a critical opportunity to harmonize policy and set priorities for species conservation and climate change adaptation. The CBD is the foundational intergovernmental agreement on biodiversity conservation and drives both government actions and donor priorities. However, the treaty itself and its existing strategic framework (the "Aichi targets") were agreed on some time ago (1992 and 2010, re-

spectively) and so need to match advances in knowledge and evidence on the immediate and devastating impacts of climate change. Over just the past few years, the frequency and severity of extreme weather events have accelerated. By one recent estimate, one-third of species may now have an increased risk of extinction from climate change (*2*).

To keep pace with changing conditions, many organisms that cannot adapt will either need to move poleward in latitude, upward in elevation, downward in water depth, or to refugial areas that might lie outside their current or historical indigenous [as defined by (*3*)] ranges. For many species, these movements are stymied by human infrastructure and disturbance. Assisted colonization—the translocation and establishment, for conservation purposes, of populations of organisms outside their historical range (*3*)—could

facilitate species conservation by moving individuals of species that cannot disperse around these barriers, allowing them to escape from shrinking climate refugia and to establish populations in new locations that have the conditions needed for population persistence. But despite having been discussed by conservation scientists for decades, assisted colonization has been deployed for climate adaptation only rarely (*4*) and often remains precluded by contradictory global policies. Private citizens in several countries, however, have been acting on their own to implement assisted colonization without guidance, oversight, or reporting (*5*). There is therefore a need for processes by which potential assisted-colonization projects could be planned, evaluated, implemented, regulated, and monitored within the framework of international treaties, intergovernmental orga-

[1]Division of Biological Sciences, University of Montana, Missoula, MT, USA. [2]Wildlife Biology Program, University of Montana, Missoula, MT, USA. [3]Wildlife Conservation Society, Bronx, NY, USA. [4]International Union for Conservation of Nature (IUCN) Species Survival Commission Conservation Translocation Specialist Group and Centre for Conservation Research, Calgary Zoological Society, Calgary, AB, Canada. [5]Department of Forestry and Environmental Conservation, Clemson University, Clemson, SC, USA. [6]Wildlife Conservation Research Unit, Oxford University, Oxford, UK. [7]Archipelago Consulting, Portland, ME, USA. [8]Department of Environmental Studies, University of New England, Biddeford, ME, USA. [9]Griffith University, Nathan, QLD, Australia. [10]IUCN Species Survival Commission, Venezuelan Institute for Scientific Investigation (IVIC) and Provita, Caracas, Venezuela. [11]Department of Environmental Science and Policy, University of California, Davis, CA, USA. [12]Department of Zoology, University of Otago, Dunedin, New Zealand. [13]School of Earth and Environmental Sciences, The University of Queensland, St Lucia, QLD, Australia. [14]Centre for Biodiversity and Conservation Science, The University of Queensland, St Lucia, QLD, Australia. [15]Wildlife Conservation Society, Global Conservation Program, Bronx, NY, USA. Email: jedediah.brodie@umontana.edu

nizations, and statutory bodies. The CBD has an opportunity to set global standards that countries can create policy on, implement, and report back on. We recommend that the CBD empower a technical committee toward creation of an assisted-colonization protocol that all countries could implement based on structured benefit-risk assessment.

UNDERSTANDING CONSERVATION VALUES, RISKS, AND GOALS

Some scientists argue that assisted colonization should never be used because of the risk of the translocated species becoming invasive; this is a major concern, because invasive species are a strong driver of global biodiversity decline (6). Others contend that the failure to embrace assisted colonization has already led to preventable extinctions (7). The debate largely hinges on differences in underlying values (e.g., about what defines "natural"), conservation objectives, assessments and perceptions of risk, and considerations of species ranges as being inherently static versus dynamic. The Convention on Wetlands of International Importance (Ramsar Convention), for example, recognizes that climate change will lead to new species entering listed wetland sites but classifies such species as alien (8). By contrast, the European Union would consider species moving on their own in response to climate change to be native (9). Indeed, "natural" and human-assisted range shifts may be critical to maintaining biodiversity and ecosystem function in our rapidly changing world (10).

Assisted colonization carries risks, but so, too, does neglecting its use. Some introduced species have catastrophic impacts on biodiversity (6). Most, though, do not decrease local diversity (11), and no documented cases of species that had undergone conservation translocation [sensu (3)] and had then become invasive are known. Whether the risks of performing assisted colonization exceed those of not performing it will vary on a case-by-case basis, meaning that assisted colonization will be warranted in some, but not all, instances. The underlying values, goals, risks, and potential benefits must be evaluated in a transparent decision framework in any assisted-colonization project (3, 12). The records from such a framework will also allow practitioners to learn from other projects, improving the efficacy of assisted colonization over time, thereby increasing the probability of success.

CLARIFYING ASSISTED-COLONIZATION POLICY

Development of clear policy guidance by a CBD-commissioned technical committee [ideally involving other organizations such as the International Union for the Conservation of Nature (IUCN)] would provide governments, scientists, and conservationists with the tools to make informed decisions. In many cases, assisted-colonization translocations may cross international borders, increasing the need for a global body to establish guidelines for best practices in decision-making. Target 3 of the current draft of the CBD Global Biodiversity Framework mandates that, "By 2030, [signatories] ensure active management actions to enable wild species of fauna and flora recovery and conservation...." Given this target, and the focus on species recovery and conservation through active management, addressing assisted colonization by the CBD is vital. Such policy development and execution can be achieved without weakening policies that aim to prevent the spread of invasive species. Although the focus here is on assisted colonization through the movement of individuals, assisted colonization could also include the movement of genetic material from one species to another beyond its historical range. Rapid advances in genome editing are increasing the feasibility of such targeted movement of genes across species boundaries as well as the ability to rewrite existing genetic sequences (13). Indeed, "assisted evolution" is already being considered for a variety of species, including corals and trees; such actions may prove critical to achieve species persistence but raise complicated questions about the identity of species and the limits humans should take to conserve them (13).

The IUCN translocation guidelines (3) do address assisted colonization but in general terms that are not operational, prescriptive, or detailed. These guidelines mention the elements that would go into risk assessments, for example, and the IUCN Conservation Translocation Specialist Group provides international training in pertinent structured decision-making, but no guidelines, protocols, or policy currently mandate such approaches in general or for assisted colonization in particular. Moreover, although there has been headway in developing assisted-colonization guidelines within some countries (12), there are no policies providing guidance regarding translocating species for assisted colonization across international borders. Cooperation between countries on assisted colonization is essential because many species will need to shift their distributions across political boundaries, necessitating cross-boundary management and monitoring through shared governance. Regulations, guidelines, and risk tolerance in regard to conservation translocations in general, and assisted colonization in particular, vary substantially across countries and taxa (12). Policy clarification is needed to support coordination among the multiple agencies in different countries that have overlapping jurisdictions over (and management objectives for) species in need of conservation action.

OPPORTUNITIES TO PROACTIVELY AVOID CLIMATE-INDUCED EXTINCTIONS

The ongoing process of negotiating the post-2020 Global Biodiversity Framework and associated goals and targets through the CBD (1) provides an opportunity to harmonize international goals and targets for species conservation and climate adaptation (see the box). Furthermore, revision of the CBD's strategic plan allows the clear distinguishing of three interacting outcomes of species movement across international boundaries: ensuring the fair and equitable sharing of biodiversity benefits, biosecurity, and movement of species to protect biodiversity values. The United Nations General Assembly is likely to adopt the CBD Post-2020 Global Biodiversity Framework into updates of its Sustainable Development Goals. The language of the CBD Global Biodiversity Framework or monitoring framework should be updated to include explicit, a priori evaluations and comparisons of the potential biodiversity benefits of conservation translocations in general and of assisted colonization in particular with the potential biodiversity costs (detriments to the species of not translocating it, along with potential impacts on other species in the recipient site; see the box). Rigorous, quantitative decision support and risk assessment tools are available to support decision-making about where and under what circumstances assisted colonization would be justified, accounting for biological and social considerations (3). The need for policy harmonization is especially acute for species that are "extinct in the wild," which have increased extinction risk but often cannot be returned to their historic ranges because conditions there are no longer suitable. Proof of this urgency was shown in November 2020, when members of the IUCN adopted a motion to improve the recovery of "extinct in the wild" species, with the support of 95% of government members voting and 99% of nongovernmental organizations voting. This included a call to "...recognize the role of populations outside historic ranges resulting from assisted colonization" (IUCN motion 119/2020).

The CBD currently mandates maintaining "viable populations of species in natural surroundings...," "...safeguarding ecosystems, [and] species...," and preventing "...the extinction and decline of known threatened species...." These will be enhanced in some cases by assisted colonization. Conversely, the CBD also mandates signatories to "...prevent the introduction of, control or eradicate those

Policy recommendations for assisted colonization of species

Focal topics on assisted colonization that we recommend for a technical committee commissioned by the Convention on Biological Diversity (CBD) to address biodiversity threats and how these topics link to specific existing concerns (nexus) of the CBD are presented.

Biodiversity threat: Climate change threatens species with increased extinction risk
CBD nexus: Zero draft goal A.2 (1) focuses on reducing extinction threat
- Define a set of criteria to support vetting of assisted colonization (e.g., which species, from where to where, when, and how)

Biodiversity threat: Moving species beyond their historic range creates an invasion risk
CBD nexus: Aichi target 9 focuses on reducing alien species and pathways to invasion
- Assisted colonization entails an act of invasion. This carries risk. Define a specific set of protocols that would be used to assess and manage this risk
- Reconsider the alien-native dichotomy in CBD language. Essl *et al.* (10) proposed describing species as "neonative" that, of their own accord, move into new areas owing to climate change. This could be expanded to apply to species moved under the auspices of assisted-colonization projects approved under international guidelines

Biodiversity threat: International translocation includes transferring potential human benefits of biodiversity across countries and cultures
CBD nexus: Zero draft goal C.1 (1) and the Nagoya Protocol (14) address sharing the benefits of genetic resource utilization equitably across countries
- Guidelines are needed to determine the appropriateness of assisted colonization in terms of affecting nations' benefits from, for example, genetic resources or tourism
- Clarify how assisted colonization relates to the Nagoya Protocol through the former's implications for societies that will lose species of historical importance or gain species that they have never encountered before

Biodiversity threat: Climate change adaptation to sustain biodiversity may include movement of genetic material among populations or species
CBD nexus: The Cartagena Protocol (15) addresses the biosecurity aspects of genetic technologies
- Directly address how movement of genetic material for the purpose of climate change adaptation fits within the Cartagena Protocol and can be accomplished with minimal threat to biodiversity
- A technical expert group should generate clear guidance over the use of genetic material in regard to preventing extinctions and increasing adaptive potential

Biodiversity threat: Stakeholder opinion can vary strongly about the benefits and risks of moving species to reduce extinction risk; introducing species into new habitats could be viewed as degrading the natural legacy of a habitat or region
CBD nexus: The CBD emphasizes mainstreaming biodiversity in socially engaged processes. The Zero draft (1) focuses on living in harmony with nature; this should include confronting issues surrounding the societal acceptability of active management tools such as assisted colonization
- Establish detailed, operational recommendations for benefit–risk assessment of potential assisted-colonization projects. This process should be socially engaged and could use existing decision support tools (5).

---

alien species which threaten ecosystems, habitats or species," which could be seen as opposing assisted colonization. This contradiction is driven by the strict use of the term "alien." The CBD should consider expanding application of the term "neonative," which was originally proposed for species colonizing new areas on their own in response to climate change (10), to species translocated through assisted colonization (see the box).

Many assisted-colonization projects are likely to involve multiple countries. Any movement of organisms across borders is regulated by the Convention on International Trade in Endangered Species of Wild Fauna and Flora if the species are included on the Convention's appendices. In such cases, permits are required from designated government authorities. Discussions could be undertaken to streamline permitting for assisted-colonization projects that are legally, scientifically, and socially justified. It would be up to each importing country to establish regulations and guidelines for assisted-colonization programs, just as they now do for any reintroduction programs. Transnational assisted-colonization projects would also need to consider poten-

tial harm in recipient countries. Again, the cross-national natures of conservation and climate change necessitate comprehensive reviews of policy and the development of policy instruments facilitating communication and shared governance.

Many governments have not established regulations or policy frameworks around assisted colonization, but the need for such efforts is urgent. The accelerating rates of climate and biodiversity change necessitate engagement from many stakeholders and sectors of society. International leadership through the CBD can provide a model for national policies. The ongoing CBD processes provide an opportunity to develop innovative conservation solutions for species that cannot move or adapt fast enough to climate change. It also creates an imperative for clear global policy guidance that facilitates the appropriate use, and inhibits the inappropriate use, of all the tools available for reducing the unacceptable, but mounting, losses of biodiversity.

The time is ripe for the global conservation community to initiate a formal evaluation of regulatory approaches for assisted colonization, along with regulatory guidance on its implementation. Ecological and social vetting and risk assessment in potential assisted-colonization projects are already likely to slow the process. Appropriate screening and analysis are essential but should not delay decision-making to the point where it is too late to achieve desired outcomes. It is essential to develop policies to confront these issues head-on so that existing conservation strategies do not become impediments through lack of forethought and planning. ∎

REFERENCES AND NOTES

1. CBD, Post-2020 documents (2020); www.cbd.int/conferences/post2020/post2020-prep-01/documents.
2. C. Román-Palacios, J. J. Wiens, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 4211 (2020).
3. IUCN, "Guidelines for reintroductions and other conservation translocations, Version 1.0" (IUCN Species Survival Commission, Switzerland, 2013).
4. N. Butt *et al.*, *Conserv. Biol*. 10.1111/cobi.13643 (2020).
5. M. W. Schwartz, T. G. Martin, *Ann. N. Y. Acad. Sci.* **1286**, 15 (2013).
6. A. Ricciardi, D. Simberloff, *Trends Ecol. Evol*. **24**, 248 (2009).
7. P. J. Seddon, *Restor. Ecol*. **18**, 796 (2010).
8. Ramsar, "Wetlands: water, life, and culture," 8th Meeting of the Conference of the Contracting Parties to the Convention of Wetlands (Ramsar, Iran, 1971), Valencia, Spain, 18 to 26 November 2002.
9. E.U. Regulation No. 1143/2014 Off., *J. Eur. Union* **57**, 35 (2014).
10. F. Essl *et al.*, *Bioscience* **69**, 908 (2019).
11. D. F. Sax, S. D. Gaines, *Trends Ecol. Evol*. **18**, 561 (2003).
12. M. W. Schwartz *et al.*, *Bioscience* **62**, 732 (2012).
13. K. H. Redford, W. M. Adams, *Strange Natures: Conservation of Nature in the Age of Synthetic Biology* (Yale Univ. Press, 2021).
14. CBD, "Nagoya Protocol" (Secretariat of the Convention on Biological Diversity, United Nations); www.cbd.int/abs/doc/protocol/nagoya-protocol-en.pdf [accessed 8 March 2021].
15. CBD, "Cartagena protocol" (Secretariat of the Convention of Biological Diversity, United Nations, 2000); https://bch.cbd.int/protocol [accessed 8 March 2021].

NANOFLUIDICS

Engineering small-ion transporter channels

Voltage-gated channels microfabricated in graphene sheets act as ion-selective transistors

By **Bruce J. Hinds**

Protein channels that span lipid membranes are the primary regulators for the transport of chemical species into and out of cells. These gates and channels enable precise chemical selectivity and markedly enhanced transport speed (*1*, *2*). Both selectivity and transport speed far exceed those of engineered membranes based on simplistic sieving and crude surface functionalization. The replication of protein-channel performance is especially challenging for mimicking potassium ion (K^+) channels with their coupled activation and selectivity gates (*3*). These channels can pump against a concentration gradient and have a 1000:1 selectivity between K^+ and sodium (Na^+) ions, despite only a ~0.38 Å difference in atomic radii. On page 501 of this issue, Xue *et al*. (*4*) created very fast ion channels between graphene two-dimensional (2D) sheets and the voltage-gating operation for an ionic transistor. Conventional microfabrication patterning and etching techniques produced lateral flow channels ~5 µm long through a stack of ~55 graphene layers.

In principle, mimicking nature's complex channels in the form of mechanically robust engineered membranes would provide a keystone for nearly perfect and arbitrary chemical separations. One major distinction between protein channels and engineered membranes is the barrier thickness. The ~4-nm-thick lipids are far too fragile compared to the ~10 µm generally needed for mechanical robustness of engineered membranes. However, these short, nanometer-scale path lengths are a critical enabler of the excellent protein-channel performance, and there are promising approaches for stabilizing channels with block copolymers and micelle stacking (*5*).

Relatively new classes of 2D materials, in sheet and tube geometries, have shown exceptional pressure-driven fluid flow that enables protein-channel–like transport rates over micrometer-thick length scales in mechanically stable materials (*6*–*8*). Near-perfect fluid slip over atomically smooth surfaces provides a generalizable approach for creating thin, chemically selective gates at the entrance of thick but fast channels that mimic protein function (*9*).

The approach of Xue *et al*. is similar to lateral graphene channels reported by Gopinadhan *et al*. that demonstrated enhanced pressure-driven fluid flow rates (*8*), but in the Xue *et al*. study, the easier-to-process reduced graphene oxide (rGO) was used, and notably, voltage was directly applied to graphene to act as a gate electrode that draws ions between the graphene sheets and opens the channels. The driving force across the membrane is a concentration gradient for a net diffusional transport process. At a critical voltage, the channel opens wide and shows unusually fast transport.

The effective diffusion coefficient is comparable to that through K^+ protein channel selectivity gates; both are two orders of magnitude faster than diffusion in bulk water. The mechanism in this graphene channel can be visualized as a capacitative and ionic analog of "Newton's cradle," where an inward swinging pendulum ball nearly instantaneously sends a ball swinging out at the opposite side of a row of hanging steel balls (see the figure). In the ion channel, capacitance creates a critical concentration of ions between graphene sheets, at which point incoming ions on the high-concentration side can force ions quickly out the other side of the membrane through concerted Coulombic motion.

The nearly atomically flat 2D structures likely helped propagate the concerted ion motion without backscattering from the

An expanding ion gate

Xue *et al*. created a graphene multilayer mimic of voltage-gated protein ion channels. A gate electrode applies a bias voltage that drives potassium ions (K^+) between the layers.

Pulling in ions
A low gate voltage starts to pull K^+ ions (red) from solution in between the graphene layers (black), stripping off their hydration shell (blue).

Electrode

Graphene multilayer substrate
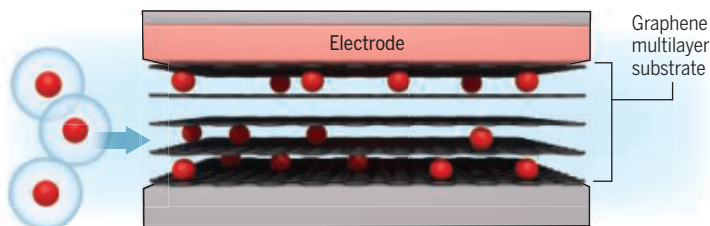
Opening the gate
Only when the layers are full at high bias and have expanded to accommodate the ions will K^+ ions flow through.
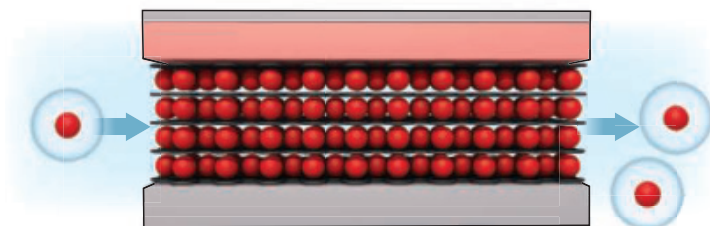
An ionic "Newton's cradle"
As K^+ ions enter the channel, they capacitively force ions out the other side, in analogy with Newton's cradle for momentum transfer.
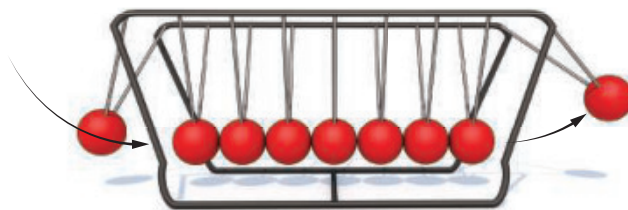
Department of Materials Science and Engineering, University of Washington, Seattle, WA 98195, USA. Email: bjhinds@uw.edu

channel surface, as this effect has not been seen in conventional or non-2D materials. A moderate ion selectivity with a K^+/Li^+ ratio of about 9 was seen and was mechanistically based on a hydration sphere diameter or ion dehydration energies to insert ions between the graphene planes. Natural K^+ ion channels are still far superior, given the 1000:1 ratio for the even more difficult K^+-Na^+ separation.

The reported fast diffusion in many ways was unexpected, because normally, a large electrode bias creates a strong attractive interaction with interfacial ions. However, in this case, the slippery 2D surfaces counterintuitively enabled very fast coupled-ion transport through Coulombic forces. Voltage gating of the channel is also an important advancement and mimics the way that many protein channels regulate cellular chemistry by acting as chemical or charge-induced valves. Most approaches to date

> "The mechanism in this graphene channel can be visualized as a capacitive and ionic analog of 'Newton's cradle'..."

have used charged ligand chemistry at pore entrances or surfaces that modulate relatively large molecular species for applications such as drug delivery (*10, 11*).

Gating small atomic ions is much more challenging, and generally only modest rectification in transmembrane currents are seen. In the work of Xue *et al.*, small ions were nearly completely blocked, and a sharp cutoff voltage was seen through a new coupled transport mechanism. Spatially addressable gates could find application in neural interfaces that require both voltage and chemical transport. ∎

REFERENCES AND NOTES

1. B. Hille, C. M. Armstrong, R. MacKinnon, *Nat. Med.* **5**, 1105 (1999).
2. K. Murata *et al.*, *Nature* **407**, 599 (2000).
3. W. Kopec, B. S. Rothberg, B. L. de Groot, *Nat. Commun.* **10**, 5366 (2019).
4. Y. Xue *et al.*, *Science* **372**, 501 (2021).
5. Y. X. Shen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 9810 (2015).
6. M. Majumder, N. Chopra, R. Andrews, B. J. Hinds, *Nature* **438**, 444 (2005).
7. J. K. Holt *et al.*, *Science* **312**, 1034 (2006).
8. K. Gopinadhan *et al.*, *Science* **363**, 145 (2019).
9. M. Majumder, X. Zhan, R. Andrews, B. J. Hinds, *Langmuir* **23**, 8624 (2007).
10. C. C. Harrell, P. Kohli, Z. Siwy, C. R. Martin, *J. Am. Chem. Soc.* **126**, 15646 (2004).
11. S. Kim, E. I. Ozalp, M. Darwish, J. A. Weldon, *Nanoscale* **10**, 20740 (2018).

ZTCG: Viruses expand the genetic alphabet

Viruses build nucleotide Z, identified in meteorites, replacing adenine in DNA genomes

By **Michael W. Grome**[1,2,3] *and* **Farren J. Isaacs**[1,2,3]

Genomic DNA is composed of four standard nucleotides, each with a different nucleobase: adenine (A), thymine (T), cytosine (C), and guanine (G). These nucleobases form the genetic alphabet, ATCG, which is conserved across all domains of life. However, in 1977, the DNA virus cyanophage S-2L was discovered with all instances of A substituted with 2-aminoadenine (Z) throughout its genome (*1, 2*), forming the genetic alphabet ZTCG. Studies revealed interesting properties of Z-substituted DNA (dZ-DNA) (*3–6*), but little of Z synthesis was understood. On pages 516 and 512 of this issue, Sleiman *et al.* (*7*) and Zhou *et al.* (*8*), respectively, characterize viral Z biosynthesis. On page 520, Pezo *et al.* (*9*) identify a Z-specific DNA polymerase that is responsible for assembling dZ-DNA from nucleotides. All three studies identify additional "Z-genomes" in diverse bacteriophages (viruses that infect bacteria), which may have offered evolutionary advantages alongside standard ATCG DNA since life began.

The nucleobases within standard DNA interact to form A:T and G:C base pairs. Two DNA single strands with complementary base sequences can recognize one another, form hydrogen bonds between matching base pairs, and zipper into a double helix. Notably, A:T forms two hydrogen bonds between bases, whereas G:C forms a stronger, more stable pair with three bonds. Within RNA, the nucleobase uracil (U) is substituted for T, forming A:U, but this does not alter the bond number. Within dZ-DNA, Z:T base pairs form with three hydrogen bonds, increasing bond number and thus stability compared with A:T (*2, 5*). This makes Z the first instance of a nonstandard nucleobase that naturally alters canonical base pairing.

Although research on dZ-DNA has been scant, three studies in the late 1980s characterized synthetic and genomic dZ-DNA, confirming three major property enhancements over standard DNA: thermal stability: dZ-DNA is more stable at higher temperatures (*5*); sequence specificity: a single dZ-DNA strand is more accurate in binding complementary DNA sequences (*4*); and nuclease resistance: dZ-DNA is resistant to degradation by nucleases that recognize and cut specific DNA sequences containing A (*3, 4*). Since those studies, the mechanical properties of helical dZ-DNA were examined (*6*), revealing increased rigidity, thermal and force stability, and a propensity to adopt a nonstandard helical form.

These features may offer evolutionary advantages in a world dominated by standard DNA. Bacteriophages (such as cyanophage S-2L) reproduce by injecting their genomic DNA into bacteria, hijacking host cellular machinery and manufacturing viral proteins to copy, build, and package new viral genomes inside the cell. To defend against infection, bacteria use a variety of mechanisms such as nucleases to destroy viral DNA. However, dZ-DNA could provide resistance to nucleases, evading host defenses. Additionally, the increased stability of dZ-DNA may permit viral persistence in extreme conditions to infect a broader range of hosts.

Despite research into dZ-DNA, little was known about the origin of Z or how it incorporates into the virus genome. Sleiman *et al.* and Zhou *et al.* used protein characterizations and metagenomic database analyses to identify key proteins of interest, determine mechanisms of action, and track sequence similarities within proteins across divergent evolutionary lineages (viruses, bacteria, and archaea). Although the entire mechanism of Z-genome replication remains unknown, two major proteins were found to be responsible for Z synthesis: PurZ (2-aminoadenylosuccinate synthetase), a virus-encoded protein that assembles a Z precursor dSMP (N6-succino-2-amino-2′-deoxyadenylate) from cellular dGMP (2′-deoxyguanosine-5′-monophosphate), and PurB (adenylosuccinate lyase), a bacterial protein that completes the synthesis of Z nucleotide dZMP (2-amino-2′-deoxyadenosine-5′-monophosphate) from dSMP (see the figure). Processing by host guanylate or nucleoside diphosphate kinases is needed to convert

[1]Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT, USA. [2]Systems Biology Institute, Yale University, West Haven, CT, USA. [3]Department of Biomedical Engineering, Yale University, New Haven, CT, USA. Email: farren.isaacs@yale.edu

dZMP into dZTP (2-mino-2′-deoxy-adenosine-5′-triphosphate) for DNA incorporation (7, 8).

Searching genomic databases for similar PurZ sequences, a diverse set of 50 to 100 putative PurZ-containing genomes were identified, predominantly from bacteriophages, but also several archaea and bacteria. Putative bacteriophage hosts vary widely in environment, geography, and phylogeny (evolutionary relation), spanning actinobacteria, cyanobacteria, and proteobacteria (7–9). Further testing confirmed two additional Z-genome bacteriophages: Vibrio phage PhiVC8 (7) and Acinetobacter phage SH-Ab 15497 (8). When examined alongside cyanophage S-2L, other major Z-associated proteins were discovered, including a putative Z-specific DNA polymerase, DpoZ, which is considered responsible for Z-genome DNA assembly. Characterization of DpoZ by Pezo et al. confirmed not only its efficiency in assembling dZ-DNA but also its capacity to discriminate between A and Z during assembly: incorporating Z while excluding A. This ability contrasts the structurally analogous bacterial DNA polymerase-I Klenow fragment that has been shown to indiscriminately incorporate both nucleotides A and Z efficiently (9). The selective nature of DpoZ may ensure exclusion of A from Z-genomes. Whether Z is incorporated into bacterial genomes is an open question worthy of exploration.

Diving deeper into phylogenetics, the viral PurZ was identified as an analog to archaeal PurA (adenylosuccinate synthetase), a protein responsible for A biosynthesis (7). This distant evolutionary connection raises questions of whether Z-specific proteins arose from bacterial host genes co-opted and adapted by viruses or are remnants of a preliminary biology, where Z-genomes existed in much greater frequency, perhaps even within cells. Phylogenetic analyses reveal that PurZ and DpoZ are often inherited together, suggesting that Z-genomes may have existed alongside canonical DNA since early evolution, before the divergence of actinobacteria, cyanobacteria, and proteobacteria ~3.5 billion years ago (9). Furthermore, a 2011 chemical analysis of a 1969 Antarctic carbonaceous meteorite revealed a composition of standard and noncanonical nucleobases of likely abiogenic extraterrestrial origin within the rock, including some deemed exceedingly rare on Earth: Z base (10). Origin-of-life researchers frequently explore meteorite compositions to gain insight into prebiotic chemical conditions on Earth, rais-

Z nucleotides in viral genomes

1 Cyanophage S-2L injects its DNA genome into its host bacterium.

2 Viral PurZ converts nucleotide dGMP into dSMP.

3 Bacterial PurB converts dSMP into dZMP.

4 dZMP is converted into dZTP by bacterial kinases.

5 DpoZ uses dZTP to replicate viral dZ-DNA to produce new virions.

6 Z:T base pairs form three hydrogen bonds, whereas A:T forms two bonds.
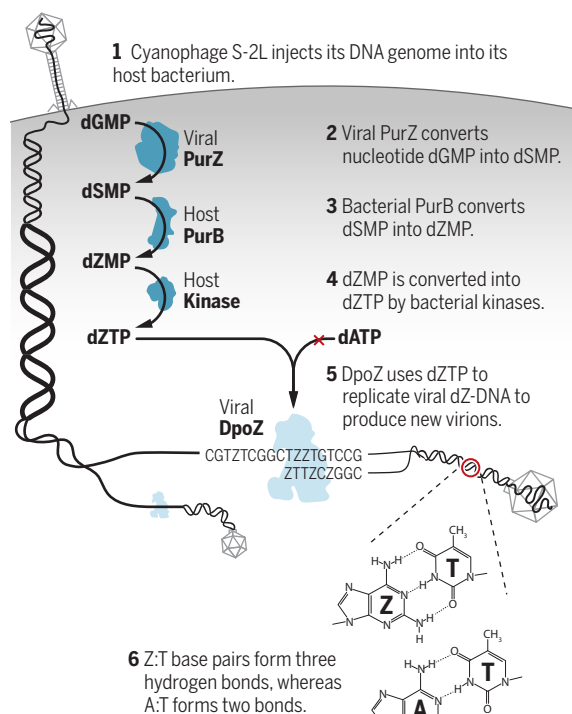
dATP, 2′-deoxyadenosine-5′-triphosphate; dGMP, 2′-deoxyguanosine-5′-monophosphate; DpoZ, Z-specific DNA polymerase; dSMP, N6-succino-2-amino-2′-deoxyadenylate; dZMP, 2-amino-2′-deoxyadenosine-5′-monophosphate; dZTP, 2-amino-2′-deoxyadenosine-5′-triphosphate; PurB, adenylosuccinate lyase; PurZ, 2-aminoadenylosuccinate synthetase.

ing a potential role for Z in early forms of life.

With capacity to synthesize dZ-DNA, researchers can harness the properties of Z directly for applications in DNA technologies. Setting the stage, a 1988 study (4) revealed that dZ-DNA probes recognize and strongly bind specific genomic sequences in certain conditions better than standard DNA. Decades later, structural DNA nanotechnology uses hundreds of synthetic DNA strands that self-assemble into three-dimensional nano-scale structures in a test tube but typically suffer from low yields and instability (11). Incorporating Z into DNA strands may offer improved thermal stability, resistance to endonucleases, structural rigidity, and greater assembly yields from enhanced sequence specificity, advancing manufacturing of personalized, injectable drug-delivery vehicles or nanoscale scaffolding for nanomachines. For synthetic biology, the enhanced properties of Z could improve the targeting capabilities and longevity of designer DNA or RNA strands used in antisense DNA and gene-targeting therapies, along with genome engineering and synthesis technologies for improved safety and efficacy (12).

Researchers have long recognized the applications of expanding genetic alphabets in cells: incorporating nonstandard bases into DNA to enhance cell manufacturing capa-

bilities, genetic isolation of strains, and biocontainment, among others (13, 14). Although researchers have incorporated noncanonical bases into DNA, they have yet to extend this technology to whole genomes. Z is a natural case that may provide insight into this endeavor. However, before knowledge of Z-genomes can be fully applied, many questions remain. For example, with the selective nature of DpoZ, a cell could theoretically build a Z-genome, but it is not known whether dZ-DNA is compatible with most DNA-interacting cellular machinery. Many proteins interact with genomic DNA to maintain a healthy, functioning cell. As Z-substitution alters DNA structure, many proteins may require reengineering to accommodate dZ-DNA.

Another question is whether Z-genomes encode Z-specific RNA polymerases, making Z-substituted RNA. Cellular translation machineries recognize RNA nucleotides to assemble proteins. If Z replaces A, this changes the code, and re-engineering of cellular machinery might be required to recognize Z and assemble protein. However, if feasible, engineered cellular machinery that recognizes Z, alongside natural machinery recognizing A, could incorporate synthetic protein subunits into natural protein production to make unnatural designer proteins for therapeutics and biomaterials or to engineer dependencies on synthetic molecules for biocontainment (15). Renewed interest in Z should spark investigations into Z-genome biology and spur new innovations in materials and biotechnology. ∎

REFERENCES AND NOTES

1. M. D. Kirnos, I. Y. Khudyakov, N. I. Alexandrushkina, B. F. Vanyushin, Nature 270, 369 (1977).
2. I. Y. Khudyakov, M. D. Kirnos, N. I. Alexandrushkina, B. F. Vanyushin, Virology 88, 8 (1978).
3. M. Szekeres, A. V. Matveyev, FEBS Lett. 222, 89 (1987).
4. A. Chollet, E. Kawashima, Nucleic Acids Res. 16, 305 (1988).
5. C. Cheong, I. Tinoco Jr., A. Chollet, Nucleic Acids Res. 16, 5115 (1988).
6. M. Cristofalo et al., Biophys. J. 116, 760 (2019).
7. D. Sleiman et al., Science 372, 516 (2021).
8. Y. Zhou et al., Science 372, 512 (2021).
9. V. Pezo et al., Science 372, 520 (2021).
10. M. P. Callahan et al., Proc. Natl. Acad. Sci. U.S.A. 108, 13995 (2011).
11. N. Seeman, H. F. Sleiman, Nat. Rev. Mater. 3, 17068 (2018).
12. W. Zhang, L. A. Mitchell, J. S. Bader, J. D. Boeke, Annu. Rev. Biochem. 89, 77 (2020).
13. D. Malyshev et al., Nature Lett 509, 385 (2014).
14. Z. Yang et al., Nucleic Acids Res. 35, 4238 (2007).
15. A. J. Rovner et al., Nature 518, 89 (2015).

MICROBIOLOGY

Gut microbiota through an evolutionary lens

Reversing some industrialization-related changes in the human gut microbiota may be risky

By **Rachel N. Carmody**[1], **Amar Sarkar**[1], **Aspen T. Reese**[2]

Although the human genome adapts on slow time scales, there is mounting evidence that industrialized lifestyles have rapidly changed the human gut microbiome (1, 2). Conceptions of health-diminishing biological incompatibility ("mismatch") arising from the disruption of human-microbe relationships negotiated over evolutionary time have led to proposals that the altered microbiota contributes to high rates of noncommunicable disease (3–6) and related calls to restore aspects of the ancestral gut microbiota through "rewilding" (4). However, appropriate applications of rewilding remain uncertain because the idea does not easily reconcile with present evidence or predictions rooted in evolutionary theory. In particular, high microbial plasticity may underpin an industrialized gut microbiota that is reasonably well adapted to the industrialized environment, even if it is then less well matched with the host. Complex tripartite human-microbiota-environment interactions present an unsolved puzzle for human health: When is it better for the gut microbiota to track versus resist environmental change?

Certain aspects of industrialized lifestyles, such as antibiotic use, have doubtless exerted strong pressure on human-microbe interactions. Although humans have always been exposed to environmental and diet-derived antimicrobial compounds, the nature and doses of antimicrobials encountered today in industrialized societies can contribute to chronic disease (6). However, although an altered microbiota may foster disease, it does not necessarily follow that health will improve upon restoring a preindustrial (ancestral) microbial state through interventions such as replacing lost gut microbial taxa, engineering microbes to perform depleted functions,

or transplanting whole gut microbial communities from donors in nonindustrial societies.

Implicit in the concept of ancestral microbiota restoration are the assumptions that the ancestral microbiota can be accurately characterized, that it promotes health, and that microbial manipulations have predictable phenotypic effects. In addition, the underlying premise that digression from evolutionarily relevant conditions compromises health assumes that natural selection elicits health, that human-microbiota mismatch has net-negative consequences, and that efforts to reestablish an ancestral microbiota in industrialized populations would reduce mismatch. The present lack of theoretical or empirical consensus on these points highlights the uncertainties involved in ancestral gut microbiota restoration.

A practical problem for restoration efforts lies in defining an ancestral microbiota. Direct assessment of historic gut microbiotas from mummies or coprolites is becoming increasingly feasible (7), but insights to date have been limited by low data quality. As an alternative, modern hunter-gatherers and other rural, nonindustrialized populations have been used as ancestral stand-ins (1, 3). However, it remains unclear whether the gut microbiotas of these living populations mimic ancestral profiles. All of the best-studied populations, including the Hadza of Tanzania, have routine access to agricultural and pharmaceutical products and are visited year-round by researchers from industrialized populations. Additionally, gut microbiotas could potentially vary between different nonindustrialized populations as much as, if not more than, they vary between industrialized and nonindustrialized populations (1). The same may have been true of ancestral microbiotas, making it difficult to define a target restoration community.

There is minimal evidence that microbial traits typically associated with nonindustrialized populations (and thus assumed to be present in an ancestral microbiota) promote health. For example, the relatively low numeric diversity of taxa and functions observed in many industrialized microbiotas has been argued to reflect a disease-associated imbalance. Although lower diversity has been observed in several disease states (8), causality remains unclear, and evidence that high diversity is beneficial is also lacking. Indeed, high diversity is theoretically predicted to be destabilizing (9). Taxon-level signatures also fail to uphold nonindustrialized profiles as uniformly healthier. For example, bifidobacteria, which are abundant in industrialized infant guts, appear to

Industrialized gut microbiota

Industrialization could affect human-microbiota-environment interactions in several ways, leading to different hypothetical effects on human health with downstream implications for rewilding the gut through ancestral microbiota restoration. Outcomes may differ based on the degree to which humans can direct (canalize) microbiota plasticity and changes in the industrialized environment to promote health.

Hypothetical "perfect" match
Area of triangle = 0
No incompatibility with respect to human health

Hypothetical ancestral state
Area of triangle = $x > 0$
Geometry unknown, shown here as equilateral

● Microbiota ● Environment ● Humans

Effects of industrialization on human health | **Implications of ancestral microbiota restoration**
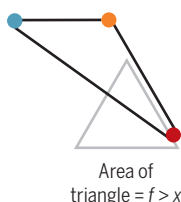
Minimal canalization

Restoration

Area of triangle = $f > x$

Area of triangle < f restoration is **beneficial**
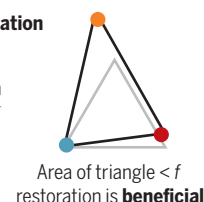
Moderate canalization

Restoration

Area of triangle = $g > x$

Area of triangle ≈ g restoration is **neutral**
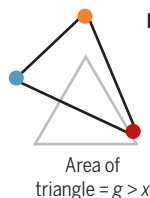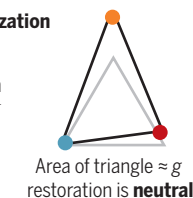
Strong canalization

Restoration

Area of triangle = $h ≈ x$

Area of triangle > h restoration is **detrimental**
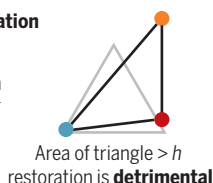
The sides of the triangle represent the degree of health-relevant incompatibility. The area of the triangle represents the overall degree of systemic mismatch.

[1]Department of Human Evolutionary Biology, Harvard University, Cambridge, MA, USA. [2]Division of Biological Sciences, Section of Ecology, Behavior, and Evolution, University of California, San Diego, La Jolla, CA, USA. Email: carmody@fas.harvard.edu; areese@ucsd.edu

be absent among infants in some nonindustrialized populations (*10*). Bifidobacteria also appear to be less abundant in nonindustrialized populations after infancy (*1*). Whether their prevalence in industrialized populations results from reduced exclusion by environmental microbes or adaptive enrichment (*10*), bifidobacteria are widely regarded as beneficial for immunity, prompting their current exploitation as probiotics.

A substantial problem facing all microbiota-targeted health interventions is that the phenotypic effects of the microbiota may be beneficial in one environment or individual but detrimental in others. For example, *Prevotella copri*, a fiber-degrader enriched in gut microbiotas from nonindustrialized populations, has shown both benefits for glucose tolerance and the propensity to exacerbate chronic inflammatory conditions, depending on context (*11*). In addition, although nonindustrialized microbiotas are generally enriched in the capacity to ferment fiber into short-chain fatty acids (SCFAs) that confer diverse metabolic and immune benefits (*3*), SCFAs can lead to context-specific developmental trade-offs with anticipated pleiotropic effects (*12*), promote weight gain through increased energy salvage, and contribute to the progression of Parkinson's disease (*13*). Similarly, two individuals could exhibit divergent responses to rewilding the gut microbiota depending on idiosyncratic factors such as parasite and viral burden, immune training, and various other gene-environment interactions.

Rewilding proposals embrace the idea that there was a time in our evolutionary past when humans were better matched to the combination of environmental and microbial conditions, but this is not necessarily the case (see the figure). Increases in both population growth and longevity with industrialization challenge this idea, and industrialization itself arose from human-directed niche construction that may be beneficial, on balance. The extent to which humans have been able to canalize (or, direct and entrain) the manifold environmental changes of industrialization to promote health remains unknown, but the burden of infectious diseases has generally decreased while the burden of noncommunicable disease has increased.

Likewise, it is unknown to what extent humans have been able to control industrialization-related changes in the microbiota. Human-microbiota interactions reflect a dynamic balance between the competing fitness interests of myriad microbial taxa and the host. Industrialization is expected to have forced gut microbes into a new state that balances fresh inputs from the host and environment while humans have presumably responded by canalizing these plastic microbial changes to the extent possible to minimize any negative consequences for fitness, as has been illustrated in many animal models (*14*). Humans certainly possess mechanisms to beneficially control the microbiota: For example, the encroachment of gut microbes into the small intestine is restricted by pH gradients, breast milk oligosaccharides shape microbial inputs to the infant immune system, and immunological responses are mounted to specific microbial products. If humans are able to exert some degree of control over changes in the microbiota, then host-microbe interactions in industrialized populations may be less detrimental than is often assumed.

Indeed, where the human capacity to canalize microbiota responses is substantial, the ability of the gut microbiota to adapt rapidly to environmental change raises the distinct possibility that, in industrialized populations, existing gut microbial profiles could allow health to a greater degree than nonindustrialized ones (see the figure). In such cases, restoring the gut microbiota to an ancestral-like state could inadvertently prove detrimental rather than beneficial.

Although ancestral gut microbiota restoration has been proposed as a possible preventative measure or treatment for noncommunicable disease (*3*–*6*), an evolutionary lens suggests fundamental challenges. For example, even in ancestral states, the capacity of humans to control the microbiota and microbiota-environment interactions to sustain health is expected to have weakened over the life course. This is because natural selection favors health only to the extent that health increases reproductive success, which declines with age. Indeed, natural selection will favor traits that exacerbate morbidity and mortality later in life if those same traits enhance fertility earlier in life, a legacy that may contribute to explaining associations between early menarche and breast cancer or between the lifetime number of viable pregnancies and metabolic disease (*15*). Thus, even if ancestral human-microbiota relationships are effectively restored, they may have limited power to ameliorate noncommunicable diseases that reach highest prevalence at older ages.

An evolutionary lens can also illuminate complementary hypotheses that advance our understanding of why human-microbe interactions have responded to industrialization as they have. For example, microbial genes for metabolizing complex carbohydrates are waning in the industrialized microbiome (*3*–*5*). This phenomenon could be viewed as an unfortunate loss of host-adapted microbes (*3*–*5*), or this loss might be viewed as metagenomic streamlining in which underutilized functions are lost while those more important in the current environment are retained. Notably, the microbes driven seasonally to undetectable levels in the Hadza were those most likely to be rare in industrialized populations, implying that ecology contributes to these differences (*1*).

Even problematic host-microbiota interactions can be appreciated as adaptive responses. For example, in the case of early-life antibiotic exposure and obesity (*6*), an evolutionary lens suggests that early-life microbiota disruption could falsely signal a volatile or resource-poor environment. Like early-life malnutrition or chronic stress, gut microbiota disruption might be expected to initiate developmental trade-offs favoring resource sparing that then contribute to obesity in resource-rich environments (*12*).

Competing fitness interests and the higher plasticity of the gut microbiota versus the human host establishes human-microbiota-environment mismatch as an omnipresent condition, both in the past and today, with variable and sometimes unpredictable effects on human health. To most effectively manipulate the gut microbiota in the service of health, the challenge is to disentangle which aspects of health are promoted by matching the microbiota more closely to the host, to the environment, or, to a lesser extent, to both. It is clear that restoration will require a scalpel, not sledgehammer, approach. Advances will be accelerated by basic research addressing the human capacity to canalize microbial and environmental change, coupled with efforts to catalog, characterize, and preserve human-associated microbes both outside industrialized contexts (*2*, *5*) and within them.

There will doubtless be scenarios where gut microbiota restoration will improve health in industrialized populations, but consideration of the sources of mismatch and their complex dynamics is necessary before pursuing targets for intervention. ∎

REFERENCES AND NOTES

1. S. A. Smits *et al.*, *Science* **357**, 802 (2017).
2. M. Groussin *et al.*, *Cell* **184**, 2053 (2021).
3. E. D. Sonnenburg, J. L. Sonnenburg, *Nat. Rev. Microbiol.* **17**, 383 (2019).
4. J. L. Sonnenburg, E. D. Sonnenburg, *Science* **366**, eaaw9255 (2019).
5. M. G. Dominguez Bello *et al.*, *Science* **362**, 33 (2018).
6. M. J. Blaser, *Science* **352**, 544 (2016).
7. C. Warinner *et al.*, *Annu. Rev. Genomics Hum. Genet.* **18**, 321 (2017).
8. C. Duvallet *et al.*, *Nat. Commun.* **8**, 1784 (2017).
9. K. Z. Coyte *et al.*, *Science* **350**, 663 (2015).
10. F. A. Ayeni *et al.*, *Cell Rep.* **23**, 3056 (2018).
11. R. E. Ley, *Nat. Rev. Gastroenterol. Hepatol.* **13**, 69 (2016).
12. I. Kimura *et al.*, *Science* **367**, eaaw8429 (2020).
13. T. R. Sampson *et al.*, *Cell* **167**, 1469 (2016).
14. K. R. Foster *et al.*, *Nature* **548**, 43 (2017).
15. G. Jasienska *et al.*, *Lancet* **390**, 510 (2017).

CONDENSED MATTER

Going with the grains

Coherent interfaces may not be vital for combining semi- and superconducting materials

By **Valla Fatemi** and **Michel H. Devoret**

Metals and insulators are primitive materials for electrical applications when compared with semiconductor materials. Semiconductors involve a massive leap in technological complexity yet are essential for data processing applications. Unsurprisingly, superconducting quantum-coherent devices have therefore been realized only with combinations of metals and insulators. More recently, proposals involving exotic states of electronic matter that would protect quantum information from decoherence (1, 2) require integration of semiconducting and superconducting properties. Aluminum (superconductor) and indium arsenide (semiconductor) have been widely investigated and exhibit a high-quality interface (3). On p. 508 of this issue, Pendharkar et al. (4) provide a proof of principle for an approach that may expand the available materials for both superconductors and semiconductors for this integration of properties (see the photo).

The history of electricity is instructive regarding the delicate nature of semiconductor materials. Metals and insulators were the basis of early applications of electricity in the 1800s because semiconductors were unreliable. Nearly a century passed before our understanding of semiconductor materials was advanced enough to gainfully apply them—most notably with the invention of the transistor in 1947, which is often associated with the dawn of the information age. Discoveries of more reliable and functional combinations of materials for semiconductor devices has been a major driver of innovation in digital computing hardware for the past 70 years.

Now, nascent superconductor-semiconductor technology proposes to merge the electric field control of low-dimensional semiconductor devices with the dissipationless nonlinear properties of superconducting devices. This combination allows the construction of, for example, the superconducting analogs of parametrically driven devices

Department of Applied Physics, Yale University, New Haven, CT 06520, USA. Email: valla.fatemi@yale.edu

for signal gain and frequency conversion. Together with a magnetic field, creating exotic physical objects such as Majorana bound states becomes possible, which may be used in qubits predicted to be inherently insensitive to a type of noise that is likely to cripple quantum computations. After some pioneering experimental results in 2012 (5), research became focused on the delicate process of fusing a semiconductor and a superconductor. A breakthrough in 2015 showed that aluminum and indium arsenide possesses a high crystalline continuity at their interface called epitaxy (3). Electrons can thus apparently coherently straddle the two materials and
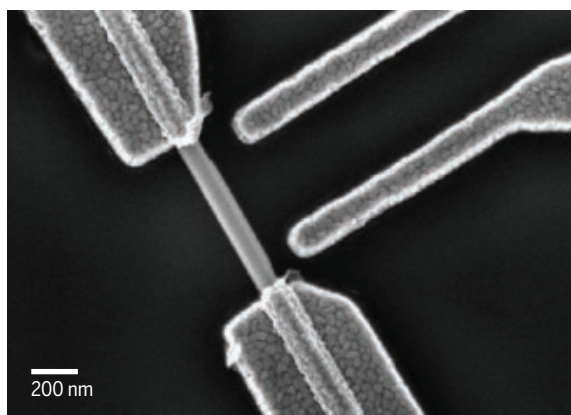
The device above was one used by Pendharkar et al. for their study of superconductor-semiconductor interfaces.

inherit both of their properties. Experiments showed that this combination produced satisfactory characteristics, which include a "hard" induced superconducting gap (6), meaning no spurious density of states at energies much lower than the characteristic energy of the superconducting state, and phenomena only possible with high electronic interface transparency (7).

Pendharkar et al.'s results question the necessity of this epitaxial matching. The superconductor of choice is tin in the usually thermally unstable beta phase; this structural phase of tin is desirable for its superconducting transition temperature, which is about four times higher that of aluminum. Through a combination of low-temperature deposition and a capping layer of a robust oxide (8), the authors found that the tin film can be largely stabilized in its targeted beta phase. The semiconductor is indium antimo-

nide, which has strong spin-orbit coupling and high electronic mobility. The authors describe several transport measurements that reveal that a hard superconducting gap is carved out of the density of states of this semiconductor (9, 10). Moreover, the hard gap is resilient to the application of a very high magnetic field, along with high interface transparencies. The interface producing these high-quality behaviors is not epitaxial; rather, it is merely locally abrupt, so that the atomic constituents of either material do not diffuse into each other. Also, the superconducting film is granular, with some grains found to be in the undesirable alpha phase. Thus, a better understanding of the relationship between material interfaces and electronic device quality is necessary to properly expand our materials inventory for superconductor-semiconductor heterojunctions.

The authors point out that this specific materials combination has relevance to the pursuit of Majorana bound states; a much broader impact is possible by using the strategy to discover new combinations. Superconductors such as tin with a higher critical temperature may help the development of dissipationless field-effect transistors for ultralow noise applications in quantum information (11). Andreev qubits also may benefit, particularly in the direction of stronger spin-orbit coupling (7, 12, 13). We can even imagine bringing these thin-film techniques to conventional superconducting qubits that might benefit from the higher gap—for example, by providing greater immunity to the poisoning effects of high-frequency photons (14). Extension of these techniques to 2D superconductor-semiconductor heterostructures or selective-area growth systems will be crucial for scalability. Although the full feasibility of these applications is understandably still an open question, an important step has been made toward developing materials for them. ∎

REFERENCES AND NOTES

1. R. M. Lutchyn, J. D. Sau, S. Das Sarma, *Phys. Rev. Lett.* **105**, 077001 (2010).
2. Y. Oreg et al., *Phys. Rev. Lett.* **105**, 177002 (2010).
3. P. Krogstrup et al., *Nat. Mater.* **14**, 400 (2015).
4. M. Pendharkar et al., *Science* **372**, 508 (2020).
5. V. Mourik et al., *Science* **336**, 1003 (2012).
6. W. Chang et al., *Nat. Nanotechnol.* **10**, 232 (2015).
7. M. Hays et al., *Phys. Rev. Lett.* **121**, 047001 (2018).
8. D. J. Carrad et al., *Adv. Mater.* **32**, 1908411 (2020).
9. I. Gu et al., *Nano Lett.* **17**, 2690 (2017).
10. S. Gazibegovic et al., *Nature* **548**, 434 (2017).
11. A. A. Clerk et al., *Rev. of Mod. Phys.* **82**, 1155 (2010).
12. L. Tosi et al., *Phys. Rev. X* **9**, 011010 (2019).
13. M. Hays et al., *Nat. Phys.* **1** (2020).
14. M. Houzet, K. Serniak, G. Catelani, M. H. Devoret, L. I. Glazman, *Phys. Rev. Lett.* **123**, 107704 (2019).

Mycobacterium abscessus in cystic fibrosis

Environmental Mycobacteria share genes and evolve to become pathogens

By **Rossa Brugha**[1,2] *and* **Helen Spencer**[1]

Mycobacterium abscessus is a nontuberculous mycobacterium (NTM) that evades the host immune system by infecting macrophages, key immune cells involved in clearing inhaled bacteria (1). In cystic fibrosis (CF) patients, M. abscessus is increasing in prevalence (2), is associated with rapidly worsening disease (3), is difficult to treat (4), and requires antibiotic regimens with substantial side effects, including severe nausea, deafness, and impaired liver function. Lung transplantation is often contraindicated in the presence of NTM and mortality is high (5). Understanding how these bacteria infect and adapt in susceptible hosts is important to develop preventive and treatment approaches. On page 483 of this issue, Bryant et al. (6) show that M. abscessus, through horizontal gene transfer, shares with other Mycobacteria a DNA methyltransferase that regulates gene expression, including those involved in antimicrobial resistance, pathogenicity, and immune evasion.

M. abscessus clones have rapidly spread among patients with CF since the last common ancestor was proposed to have emerged around 1978 (7). Bryant et al. show

[1]Respiratory Medicine and Cardiothoracic Transplantation, Great Ormond Street Hospital, London, UK. [2]Infection, Immunity and Inflammation, UCL Great Ormond Street Institute of Child Health, London, UK. Email: helen.spencer@gosh.nhs.uk

the adaptations that M. abscessus undergoes once infection has been established, including a change in morphotype (from smooth to rough, which is associated with greater antibiotic resistance and more invasive behavior), and ribosomal RNA adaptations that confer antibiotic resistance, which were associated with a decreased likelihood of indirect patient-to-patient transmission. These adaptations resulted in the generation of multiple series of subclones within the same patient, a finding documented to occur in Pseudomonas aeruginosa infections in CF patients (8). Like M. abscessus, P. aeruginosa infection in CF is associated with more severe disease, including reduced lung function and increased hospital admissions. If M. abscessus can evade host defenses and develop antibiotic resistance in a similar manner to P. aeruginosa, this poses a serious treatment challenge because the antibiotic susceptibility data available from airway sample cultures are likely to reflect only a portion of the overall NTM population in the lung (9). Treatments selected on the basis of the sensitivities of one group of bacteria in one lobe of a lung may be ineffective or drive the development of antibiotic resistance in a separate population in a different area of lung tissue.

The extent to which between-patient spread of NTM occurs is highly contentious. Multiple environmental sources of this ubiquitous organism may confound analyses attempting to elucidate causation once a new infection is discovered. Outbreaks

of NTM have occurred in health care settings associated with contaminated water supplies, particularly in new buildings, and colonized heater-cooler circuits used in cardiopulmonary bypass (10). The mechanisms of direct patient-to-patient spread are less clear, with fomites (clusters of bacteria contaminating surfaces that can transmit infection by contact or aerosolization) being implicated in an outbreak in adults with CF in Seattle, WA (11). Multicenter studies of transmission have shown markedly different results; a 16-year dataset using a whole-genome sequencing (WGS) approach covering four hospitals in two countries demonstrated only one possible patient-to-patient transmission event (in siblings) (12). Conversely, a separate retrospective report from 11 centers in six countries, using similar parameters for determining clusters, concluded that the majority of M. abscessus infections were acquired through person-to-person transmission rather than indirectly from environmental sources (7). Bryant et al. report cross-infection between patients, with additional analyses suggesting this occurs preferentially in M. abscessus bacteria that have not yet undergone the evolutionary changes necessary to establish infection in the CF lung (see the figure).

What are the implications for clinical care? International guidelines (13) recommend annual surveillance for NTM species in patients with CF, and a single isolate of NTM does not necessarily mean an immediate recourse to treatment. The reasons for this are myriad, based on the difficulty in culturing NTM from cough swabs in non–sputum-producing patients, making surveillance challenging. Additionally, there are concerns about laboratory contamination of samples and apparent spontaneous clearance of NTM: In a recent report, 7 of 34 (21%) children had M. abscessus on one occasion, which then cleared spontaneously (4). The guideline treatment approach to M. abscessus is a 3- to 12-week course of intravenous antibiotics followed by prolonged tri-

Environmental sources and evolution of *Mycobacterium abscessus*

The bacterium *Mycobacterium abscessus* has numerous possible environmental sources, which can infect susceptible hosts through the lungs, wounds, or catheters. Once the host is infected, *M. abscessus* can evolve through gene transfer, creating subpopulations that are more resistant to antibiotics and evade the host immune response.

Sources
Amoeba
Fomites
Shower heads

Hosts
Lungs:
• Cystic fibrosis
• Bronchiectasis
• Age
• Immunosuppression
• Smoking and vaping
Wounds
Catheters

Evolution
Gene transfer → Antibiotic resistance → Multiple lung subpopulations → Multiple different antibiotic susceptibilities

GRAPHIC: N. CARY/SCIENCE
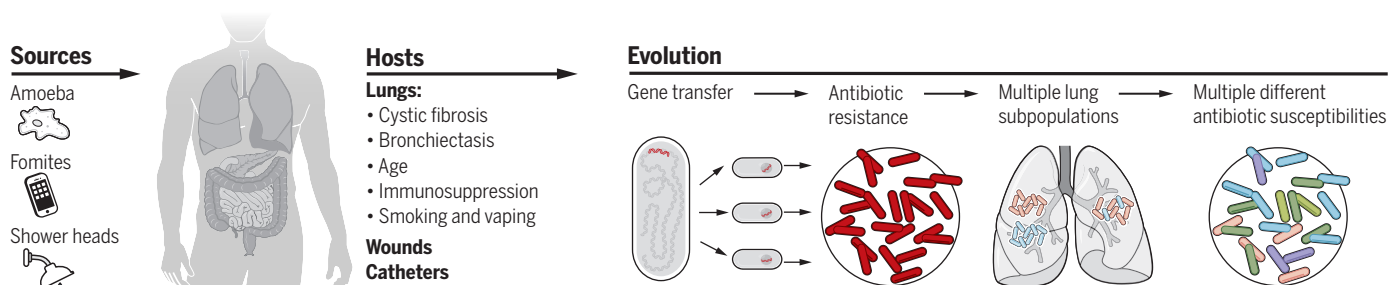
ple or quadruple antibiotic therapy through oral and inhaled routes, for at least a year, depending on subsequent culture results (*13*). Initiation of an unpleasant 12-month treatment course in a setting where a substantial proportion of patients, in retrospect, do not need the therapy is difficult to justify. But the data from Bryant *et al*. suggests that there may also be a risk in waiting. If it is indeed the case that *M. abscessus* evolves within the lung to evade the immune system and antibiotics, timely treatment is more likely to result in successful eradication of the infection.

A WGS approach, using highly accurate phenotyping of new isolates of NTM, may help to risk-stratify treatment decisions, and a carefully designed randomized clinical trial of current versus more aggressive approaches to NTM infection in this setting, including early treatment decisions and newer antibiotics, is required. Efforts to inform this trial design are ongoing (NCT02419989), but barriers remain, including availability of repurposed tuberculosis drugs that target *M. abscessus* (bedaquiline and clofazimine) as well as newer agents (such as relebactam, omadacycline, and tedizolid). Whether therapies such as specific bacteriophages (viruses that target bacteria) can be used as adjuncts (*14*), or in place of antibiotics, remains to be determined.

In the 30 years since the identification of mutations in the cystic fibrosis transmembrane regulator (*CFTR*) gene on chromosome 7 as the genetic basis of CF, care has changed markedly. For the majority of patients with the health care resources to support them, new "modulator" therapies (such as elexacaftor, ivacaftor, tezacaftor) promise improved lung health and extension of life spans with CF (*15*). The data presented by Bryant *et al*. are a timely reminder that individuals with CF remain at risk of infection with this highly resistant organism, and that vigilance regarding infection control measures must remain a key focus in CF care. ∎

REFERENCES AND NOTES

1. M. D. Johansen, J. L. Herrmann, L. Kremer, *Nat. Rev. Microbiol.* **18**, 392 (2020).
2. N. Z. Abidin *et al*., *J. Cyst. Fibros.* 10.1016/j.jcf.2020.09.007 (2020).
3. T. Qvist *et al*., *J. Cyst. Fibros.* **15**, 380 (2016).
4. D. A. Hughes, I. Bokobza, S. B. Carr, *Eur. Respir. J.* 10.1183/13993003.03636-2020 (2021).
5. E. Kavaliunaite *et al*., *Transpl. Infect. Dis.* **22**, e13274 (2020).
6. J. M. Bryant *et al*., *Science* **372**, eabb8699 (2021).
7. J. M. Bryant *et al*., *Science* **354**, 751 (2016).
8. S. T. Clark, D. S. Guttman, D. M. Hwang, *FEMS Microbiol. Lett.* **365**, fny026 (2018).
9. L. P. Shaw *et al*., *Clin. Infect. Dis.* **69**, 1678 (2019).
10. J. A. Gilbert, *Lancet Respir. Med.* **5**, 384 (2017).
11. M. L. Aitken *et al*., *Am. J. Respir. Crit. Care Med.* **185**, 231 (2012).
12. R. M. Doyle *et al*., *Clin. Infect. Dis.* **70**, 1855 (2020).
13. R. A. Floto *et al*., *Thorax* **71** (suppl. 1), i1 (2016).
14. R. M. Dedrick *et al*., *Nat. Med.* **25**, 730 (2019).
15. P. G. Middleton *et al*.; *N. Engl. J. Med.* **381**, 1809 (2019).

CORONAVIRUS

How SARS-CoV-2 first adapted in humans

An early spike protein mutation promotes transmission and will shape the next vaccines

By **Hyeryun Choe** *and* **Michael Farzan**

Viruses need entry proteins to penetrate the cells where they will replicate. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) version is called the spike or S protein. The S protein, also the target of the current vaccines, is quickly adapting to its new human hosts. It took its first major step in this direction early in 2020, when its amino acid 614 (of 1297) changed from an aspartic acid (D) to a glycine (G). Viruses bearing this D614G mutation transmit among humans more rapidly and now form the majority in circulation. On page 525 of this issue, Zhang *et al*. (*1*) use careful structural analyses to reveal how D614G changed the S protein to accelerate the pandemic.

Early in the pandemic, in the scramble to create tools to study SARS-CoV-2, investigators developed pseudovirus systems to measure infection in a safe, easily quantifiable way. These systems express a viral entry protein on the surface of a reporter virus used to monitor cell entry and have been used for years to study many such proteins, including the S protein of "classic" SARS-CoV-1. Frustratingly, pseudoviruses built from the SARS-CoV-2 S protein produced much lower signals than those based on the very similar SARS-CoV-1 S protein. This was perplexing because biochemical studies of SARS-CoV-1 and SARS-CoV-2 S-protein receptor binding domains (RBDs) made clear that the SARS-CoV-2 RBD bound their common receptor, angiotensin-converting enzyme 2 (ACE2), with higher affinity than that of SARS-CoV-1 (*2*, *3*). Faced with inefficient assays, many virologists landed on the same solution as their structural biology colleagues: Mutate the S-protein site that is cleaved by furin-like proteases in virus-producing cells (*2*). This change kept the S-protein S1 domain, which contains the RBD and binds ACE2, covalently linked to its S2 domain, which anchors the S protein to the virion. Notably, some—but not all—of these furin-site mutations significantly improved pseudovirus infection of cells (*4*).

This fix solved a technical problem, but it deepened a mystery. Although a number of distantly related coronaviruses carry furin cleavage sites at their S1-S2 boundaries, the SARS-CoV-1 S protein, and those of all known bat-derived viruses from the same Sarbecovirus lineage, lack this site. Instead of being cleaved in virus-producing cells, their S proteins are cleaved by different proteases while the virus is engaging ACE2 in the next, yet-to-be-infected cell (*5*). As it happened, furin-site mutations that improved SARS-CoV-2 S-protein function in pseudoviruses allowed the modified S protein to work with these later-stage enzymes, just like the SARS-CoV-1 version. Why then did the SARS-CoV-2 furin site persist, even though it made infection in cell culture less efficient? Indeed, viruses passaged in culture regularly lost this site. Does it somehow improve viral transmission? Would it eventually disappear over the course of the pandemic?

In the summer of 2020, Korber *et al*. sounded an alarm about a "mutation of concern," namely D614G (*6*). In the laboratory, this change obviated the need to eliminate the S-protein furin site, apparently correcting a design flaw associated with this unusual cleavage site (*4*, *6*). Animal studies with otherwise identical viruses showed markedly greater replication of the D614G variant in the upper respiratory tract, a site important for transmission (*7*, *8*). By contrast, no significant differences between the two viruses were seen in the lower respiratory tract, a site responsible for more severe disease (*7*). These observations are consistent with the current consensus that D614G, now present in most circulating viruses, enhances viral transmission, but unlike more recent acquired mutations in S1 [e.g., Asn^{501}→Tyr (N501Y)], it does not change the rate of hospitalizations.

The underlying mechanism for this fitness advantage remained a point of controversy. Here, a second unusual property of the S protein, in this case shared with SARS-CoV-1, became relevant. The SARS-CoV-2 S protein, like most entry proteins of viruses with a lipid membrane, assembles into trimers. Typically, during the process of virion assembly, viral entry proteins subtly change their

Department of Immunology and Microbiology, Scripps Research Institute, Jupiter, FL, USA.
Email: hchoe@scripps.edu; mfarzan@scripps.edu

conformations, but it is unusual for these proteins to break their three-fold symmetry before they bind their receptor. However, the mature SARS-CoV-2 S protein often assumes an asymmetrical arrangement whereby one of its three RBDs assumes an open or "up" conformation [1, 9]. Only RBDs in this up conformation can bind ACE2. Once it does so, the S1 domains dissociate from S2, and S2 undergoes a pronounced rearrangement to a "postfusion" state. The released energy of this rearrangement drives the viral and cell membranes to fuse and gives the virus access to the cell interior.

To explain D614G fitness, some investigators focused on the impact of D614G on the frequency with which this one-up conforma-

bridge is initially counterintuitive because it would loosen the association between S1 and S2, although it might ease the movement of the RBD into the up configuration. However, the structures from Zhang *et al.* show that a major difference between S proteins with and without D614G is the visibly greater ordering in G614 S proteins of a region spanning residues 620 to 640, which the authors call the "630 loop." This loop is just downstream of G614. It is therefore possible that either the loss of the D614-K854 salt bridge, or the greater backbone flexibility that a glycine affords, helps the 630 loop nestle more tightly in a canyon formed by two larger S-protein domains (the amino-terminal domain and carboxyl-terminal domain 1). Regardless, this

because it prevents premature S1 shedding (see the figure).

These structural studies have real-life implications. All current vaccines are based on the original, unstable D614 form of the S protein [13]. Fortunately, most vaccine developers, including Moderna and Pfizer-BioNTech, took a lesson from studies of SARS-CoV-1 and Middle East respiratory syndrome (MERS) coronavirus to slow S-protein shedding by introducing non-native prolines into S2 [14]. Those who developed the Johnson and Johnson and Novavax vaccines had the prescience to also remove the furin site. By contrast, the developers of the University of Oxford–AstraZeneca vaccine opted for the wild-type S protein (containing D614), as is also the case for the inactivated virus vaccine produced by Sinovac. To be clear, other variables, especially antigen-delivery systems, likely account for efficacy differences among these vaccines. However, apples-to-apples studies in animals make clear that both engineered prolines and furin-site ablation contribute to vaccine effectiveness [15]. It is almost certain that the next round of vaccines, those better reflecting the S-protein variants now in circulation, will include D614G. Vaccines that express unmodified S proteins with G614 may enjoy a relative jump in potency because this change compensates for the lack of engineered stabilizing mutations.

The work of Zhang *et al.* also reveals more about the natural history of the virus. The notable emergence of D614G suggests that the acquisition of a destabilizing furin site was a recent event. The virus could easily lose this site, as it does frequently in cell culture systems, implying that it in some way facilitates human transmission. This is not a conclusion that most students of human coronaviruses would have anticipated, given that SARS-CoV-1, which transmits with reasonable efficiency, lacks this site, whereas the more distantly related MERS coronavirus bears this site and transmits poorly. How the SARS-CoV-2 furin site promotes new human infections remains a key open question in the field. ∎

Enhancing viral transmission

The Gly[614] (G614) mutation in spike (S) increases ordering of the 630 loop compared with wild-type Asp[614] (D614). This prevents the premature S1 shedding often seen with wild-type S proteins, ensuring that more S protein remains in a fusion-ready "one-up" state, with one receptor binding domain (RBD) exposed within the trimer, ready to bind angiotensin-converting enzyme 2 (ACE2) on host cells, increasing infection efficiency.
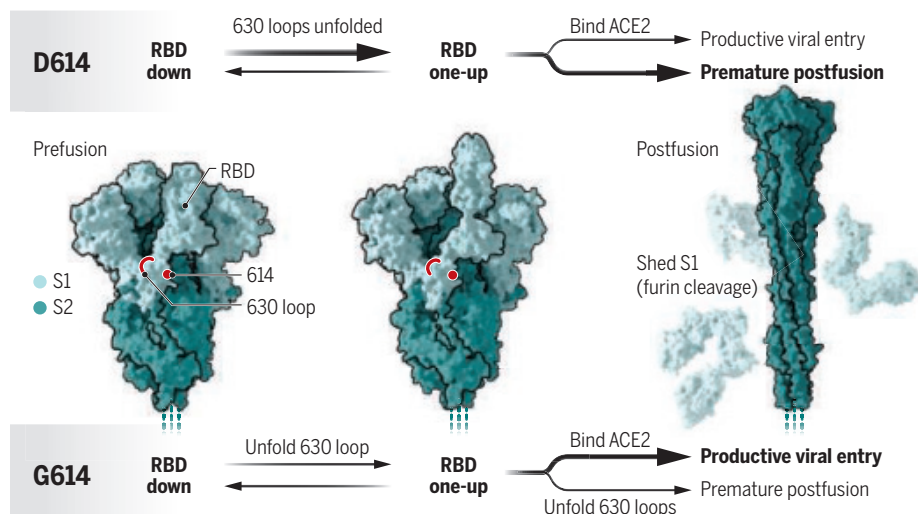
tion could be found, suggesting that more efficient engagement of the receptor accounted for the enhanced transmissibility of viruses bearing this mutation [10, 11]. Others noted that S proteins of D614G-expressing viruses fell apart less frequently, an effect perhaps amplified in the challenging environment of a living organism. They observed that D614G helped the S1 domain cling to S2, preventing S2 from prematurely and unproductively assuming its postfusion conformation [4, 9, 12]. Thus, the virus had more functional S proteins that could bind and infect the next cell.

To cut through this controversy, Zhang *et al.* solved the structure and provided detailed analyses of both D614 and G614 S proteins in multiple states. They first noted that, as they and others had previously observed, the loss of D614 in S1 breaks an ionic bond to a proximal lysine, K854, in S2 [9]. Loss of this salt

loop is found in a more rigid and stable arrangement between these domains when residue 614 is a G than when it is a D.

The key is that both the RBD-up conformation and dissociation of S1 from S2—enabled by furin cleavage—require disordering of the 630 loop. Thus, the RBD-up conformation can be more easily accessed with the original D614 S protein, but once this conformation is achieved, this S protein is more likely to fall apart entirely owing to premature shedding of its S1 domain. Conversely, with G614, more energy is required to achieve a one RBD–up state, but dissociation of S1 from S2 also becomes less favorable because the remaining folded 630 loops continue to hold the trimer together. Thus, the D614G variants have more S proteins in the up orientation because the next, irreversible step toward inactivation is slower. Infection with D614G is more efficient

REFERENCES AND NOTES

1. J. Zhang *et al.*, *Science* **372**, 525 (2021).
2. A. C. Walls *et al.*, *Cell* **183**, 1735 (2020).
3. W. Li *et al.*, *Nature* **426**, 450 (2003).
4. L. Zhang *et al.*, *Nat. Commun.* **11**, 6013 (2020).
5. M. Hoffmann *et al.*, *Cell* **181**, 271 (2020).
6. B. Korber *et al.*, *Cell* **182**, 812 (2020).
7. J. A. Plante *et al.*, *Nature* **592**, 116 (2021).
8. B. Zhou *et al.*, *Nature* **592**, 122 (2021).
9. Y. Cai *et al.*, *Science* **369**, 1586 (2020).
10. D. Weissman *et al.*, *Cell Host Microbe* **29**, 23 (2021).
11. L. Yurkovetskiy *et al.*, *Cell* **183**, 739 (2020).
12. J. Juraszek *et al.*, *Nat. Commun.* **12**, 244 (2021).
13. J. Mellet, M. S. Pepper, *Vaccines* **9**, 39 (2021).
14. C.-L. Hsieh *et al.*, *Science* **369**, 1501 (2020).
15. J. Yu *et al.*, *Science* **369**, 806 (2020).

RETROSPECTIVE

David William Schindler (1940–2021)

Trailblazing aquatic researcher, advocate, and whole-ecosystem experimentalist

By **Vincent L. St. Louis[1], Sherry L. Schiff[2], Carol A. Kelly[3]**

David William Schindler, innovative ecologist, died on 4 March at the age of 80. Dave documented humanity's impact on global freshwaters with seminal research on eutrophication, acid rain, climate change, organic contaminants, and oil sands mining. He then used his formidable energy and dedication to advocate for regulation of harmful industrial practices and to educate the public about why these regulations were needed.

Dave was born in Fargo, North Dakota, on 3 August 1940 but raised near Barnesville, Minnesota. He enrolled in an engineering physics program at North Dakota State University in 1958, but a chance reading of Charles Elton's book *The Ecology of Invasions by Animals and Plants* during a summer biology internship changed his life forever. Switching majors, he earned his bachelor's degree in zoology in 1962. He then pursued his PhD in ecology with Elton as a Rhodes scholar at the University of Oxford, studying energy relations in aquatic food chains. Just 2 years after receiving his doctorate in 1966, he was scouted to become the first scientific director of the Experimental Lakes Area (ELA), a Canadian field station newly created for the purpose of using whole-ecosystem manipulations to address freshwater problems of international concern. In 1989, he left to become the Killam Memorial Chair and Professor of Ecology at the University of Alberta, a position he held until 2013. Thereafter, although nominally retired, Dave continued uninterrupted in his scientific research and advocacy. He was inspired by his wife, ecologist Suzanne Bayley, and his three children.

Dave's singular success arose from his early recognition that complex ecosystems function in ways not predictable from smaller bottle and aquarium experiments and that the whole system therefore needs to serve as the unit of experimentation. When tasked with bringing this approach to life at the ELA, a set of 58 headwater lakes in northern Ontario, he intuitively assembled a team with expertise in hydrology, lake physics, biogeochemistry, microbiology, plankton, benthic invertebrates, and both small and large fish. He established a year-round boot camp for scientists to do one-of-a-kind collaborative freshwater research. Science was discussed at communal meals, in boats, in laboratories, and on the swimming dock. As Dave was a habitual early riser, it was an unfortunate scientist who did not bring their best ideas to breakfast.

By changing only one factor at a time, Dave was able to study the cascading effects of that change through different parts of the lake ecosystem and, crucially, to elucidate the underlying mechanisms of change. Early whole-lake experiments demonstrated that phosphorus was the limiting nutrient for photosynthesis in lakes (nitrogen and

carbon could easily be obtained from the atmosphere) and that limiting phosphorus inputs would therefore solve eutrophication. By adding sulfuric or nitric acids to lakes, he showed that small decreases in pH alone could disrupt food chains, leading to starvation of fish. These efforts resulted in two of the most iconic photographs in freshwater research. One photo showed that the side of a divided two-basin lake receiving experimental phosphorus additions turned pea green from an algal bloom. The second photo showed a gaunt lake trout from an experimentally acidified lake. Dave used these photos to rally public support and force needed legislation.

While at the ELA, Dave also fortuitously began amassing the world's largest dataset on climate change impacts on freshwaters through the basic monitoring of unmanipulated reference watersheds, which has continued now for more than 50 years as Earth has warmed. An initial publication after only 20 years provided an early preview of the effects of warming on boreal lakes.

The Alberta oil sands industry claimed that petroleum-related compounds in the Athabasca River were all natural, but Dave was instinctively skeptical. His ensuing studies showed that the industry was a regional source of organic contaminants and metals. At a news conference, Dave showcased fish with unusual tumors downriver of the oil sands developments, and this kind of easily understood public demonstration led to the establishment of an independent inquiry.

Unfettered by societal norms, Dave freely followed his innate scientific intuition. He had no patience for half-baked ideas, political wrangling, pessimistic whining, or experimental roadblocks. At the same time, Dave was a strong supporter of fledgling scientists, empowering even undergraduate summer students to contribute as equals. He taught us that ruffling feathers probably meant you were actually getting somewhere, and he reassured us with his pithy maxim that "Science bats last," meaning that scientific evidence will always prevail.

Dave had energy to burn. An avid woodworker, canoer, hunter, fisherman, and mountain biker, he was always happiest outdoors. He and Suzanne also competitively raced the multitudes of champion sled dogs they raised.

A prodigious communicator of his scientific research, Dave published more than 350 papers. He received upward of 100 prestigious awards, including the first Stockholm Water Prize, and was made an Officer of the Order of Canada. Each year, he gave dozens of presentations, many of them public talks, and unscripted interviews on television and radio. He was also a vocal champion for Canadian Indigenous communities affected by the broader population's impacts on their treaty lands and traditional freshwater resources.

To many, Dave seemed larger than life. His research fundamentally changed freshwater science and the lives of generations of students. We each had the good fortune of having Dave as a mentor, colleague, and friend for more than 40 years. Dave's scientific and public legacy remains as both an inspiration and a call to action. ■

[1]Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2E9, Canada. [2]Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, ON N2L 3G1, Canada. [3]R&K Research Inc., Salt Spring Island, BC V8K 2J3, Canada. Email: vince.stlouis@ualberta.ca

PHOTO: JOHN ULAN

BOOKS *et al.*

SPACE COLONIZATION

The biologically optimized spacefarer

A geneticist proposes a radical plan to help humanity live comfortably on other planets

By **Konrad Szocik**

Humans have the capacity for very long-term thinking, yet we are often careless about the welfare of Earth and reckless when it comes to our destructive abilities. As such, we may eventually be confronted with the necessity of colonizing the Universe. In such a scenario, might it be possible to genetically modify humanity to optimize ourselves for the new environmental challenges we will face on other planets? In his new book, *The Next 500 Years*, geneticist Christopher Mason argues, "It is no longer a question of 'if' we can engineer life—only 'how'" and "Engineering is humanity's innate duty, needed to ensure the survival of life."

In the book's first chapter, Mason discusses his role in the NASA Twins Study (*1*), which sought to measure various physiological markers in astronaut Scott Kelly during a year aboard the International Space Station and compare these indices with comparable data from his identical twin, astronaut Mark Kelly, who remained earthbound during the same period. The study, he argues, which found (among other things) that the expres-

sion of a number of genes related to DNA repair remained elevated 6 months after Scott returned to Earth, has enormous implications for understanding exactly how the space environment affects the human body.

Mason begins the next chapter by expressing an almost messianic belief in what he sees as the uniquely human responsibility of preserving life, not only our own but that of all of the other species in the Universe. He fails to consider, however, numerous ethical and logical possibilities, including, for example, our obligations to future humans as well as the potential cost of attempting to preserve life above all other considerations.

Over the next 10 chapters, Mason discusses the peculiarities, challenges, and targets of each of the 10 phases into which he divides his proposed space colonization project. The first phase, already completed between 2010 and 2020, is a detailed study of the human genome. The second, which he proposes will be completed over the next 20 years, is dedicated to preliminary engineering of genomes. One promising example discussed here is the successful integration into human cells of a gene known to promote protection from radiation in tardigrades, microscopic creatures noted for their extraordinary resilience.

The next stage in Mason's proposed project

> Mason's plan culminates in 2500, by which time he predicts we will be able to direct our own evolution.

involves a series of long-term trials of human and cellular engineering, which he anticipates could be completed between 2041 and 2100, that focus on enhancing our genomic defenses to space radiation. This is the period, he writes, in which "all the genes, cells, and even potentially organs of any organism can become a component of a human cell."

Between 2101 and 2150, Mason suggests that we begin preparing humans for space. This, he argues, will be a period in which humans will no longer be beholden to naturally transmitted genomes. Genetic modification will be widely used, he predicts, "and it is likely that a significant proportion of people in the United States will be zygotically edited or will be the product of someone who is."

In phase 5 (2151–2200), Mason predicts that people will be able to move fairly freely between Earth and space bases and that genetic modification will be available to those who want it. In phase 6 (2201–2250), we will make humans tolerant of increasingly more extreme conditions. And in phase 7 (2251–2350), Mason thinks that people will be living in fully developed Martian colonies, intergenerational interstellar travel may be possible, and the DNA of newly discovered life-forms could potentially be sequenced and used to further refine our genomes.

After phase 8 (exoplanetary settlement; 2351–2400), we will enter phase 9 (2401–2500), an era of nearly limitless human possibility. "Humans have the ability to control their underlying genetic code, controlling for how their molecules fundamentally change in response to stimuli and enabling new abilities," writes Mason. "This will enable an

The Next 500 Years
Christopher E. Mason
MIT Press,
2021. 296 pp.

unprecedented ability to build, edit, and transplant cross-kingdom combinations of genomes, which we will need to survive on new worlds." In Mason's final phase, "humans will have mastered an ability that truly sets us apart from other species— the ability to direct our own, and other organisms', evolution."

Despite sharing a sympathy for the idea of radical genetic modification of humans, I disagree with Mason's fundamental premise that it is, above all, our duty to protect and extend life in the Universe. There must be a quality to the life we defend, and I question whether radically modified humans in remote space colonies will be truly happy that such a decision has been made on their behalf. ∎

The reviewer is at the Department of Social Sciences, University of Information Technology and Management in Rzeszów, 35-225 Rzeszów, Poland. Email: kszocik@wsiz.edu.pl

REFERENCES AND NOTES

1. F. E. Garrett-Bakelman *et al.*, *Science* **364**, eaau8650 (2019).

IMAGE: MERLIN74/SHUTTERSTOCK

NEUROSCIENCE

Unlocking the secrets of self-awareness

Long neglected by researchers, metacognition has made its way to the scientific spotlight

By **Christian C. Ruff**

For millennia, religious thinkers and philosophers have cited humanity's self-awareness—that is, our ability to think about our own mind and character—as being key to the uniqueness of our species. Carl Linnaeus's groundbreaking biological taxonomy (*1*) likewise characterized our genus by the words "Homo. Nosce te ipsum" ("Man. Know thyself").

Given our long-standing interest in self-awareness, it is surprising how little science has traditionally had to say about it. What features of our brains enable us to think about ourselves? What are our strengths and weaknesses in this respect and how do they influence how we decide, learn, and interact? Can we train self-awareness, and how does this improve our performance? In the past three decades, however, research addressing such questions has been picking up speed. In *Know Thyself*, cognitive neuroscientist Stephen Fleming synthesizes this multifaceted research into an admirably coherent narrative and outlines how the resulting knowledge may be applied to solve societal problems.

Writing about self-awareness is challenging because concepts such as "self" and "awareness"—let alone the combination thereof—are hard to define. The book does not get lost in this epistemological Bermuda triangle but rather conceptualizes self-awareness as the set of mental and brain processes that keep track of our percepts, thoughts, and actions.

Not all of these metacognitive processes concern the self in a philosophical sense, Fleming notes, and not all of them need to be conscious. A helpful metaphor in the book compares the human brain to a flying plane that is largely guided by autopilot technology but can be flexibly controlled by the pilot whenever needed. For our behavior, the autopilot is the unconscious,

"implicit" part of metacognition, and the pilot is the "explicit" metacognition that we can consciously report.

Fleming begins by summarizing the psychology and neuroscience of these metacognitive processes. Implicit metacognition, he notes, is evident in many seemingly low-level brain processes, ranging from the sensory brain cells that signal the uncertainty associated with particular percepts, to brain cells that activate when we commit action errors (think: mistyping on a keyboard). All of these implicit signals can be read out in the service of explicit metacognition, when, for example, we need to judge our confidence in having

Self-awareness operates on a need-to-know basis, taking a back seat during well-learned tasks.

chosen the right action. This latter ability depends on specific brain areas in the prefrontal cortex and is independent of the basic perceptual and motor abilities it serves to monitor.

Explicit metacognition, meanwhile, depends on our ability to think about the mental states of others—an ironic twist nicely summarized by the caption of a cartoon that appears in the book: "Of course I care about how you imagined I thought you perceived I wanted you to feel."

It is eye-opening to realize how many fields of human endeavor depend not just on our skills and knowledge but also on our ability to estimate our competence.

**Know Thyself:
The Science
of Self-Awareness**
Stephen M. Fleming
Basic Books, 2021. 304 pp.
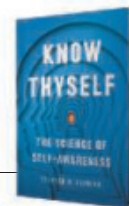
Obvious examples can be found in education, politics, the legal system, corporate decision-making and leadership, news and social media, and, indeed, any collaboration in which people pool their expertise. The book illustrates the role of metacognition in these diverse fields with elegant combinations of philosophical considerations, basic science findings, and more applied examples.

Fleming even ventures into the near future, sketching how artificial intelligence with superhuman computational abilities but no self-awareness may become disconnected from humanity at best and outright catastrophic at worst. Emerging ideas on how we may address this problem—for example, by endowing intelligent machines with coarse self-awareness or by ensuring that self-aware humans remain at the helm—only serve to prove how little we have appreciated our own prodigious metacognitive abilities.

In the end, the book makes a convincing case that self-awareness is a key feature of human existence and that our growing knowledge about it will be important for addressing many of our societal problems. One may quibble that the book somewhat understates this point, because it focuses on metacognition and does not cover our ability to monitor our emotions, another key aspect of self-awareness that has major implications for health and well-being. However, the literature on this topic is so diverse that doing it justice would likely require several additional volumes. As it stands, Fleming's book finally heaves metacognition into a long-deserved place in the scientific spotlight. ∎

The reviewer is at the Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Zurich, Switzerland. Email: christian.ruff@econ.uzh.ch

REFERENCES AND NOTES
1. C. Linnaeus, *Systema Naturae* (Haak, 1735).

PHOTO: ISTOCK.COM/CHABYBUCKO

2021 Mike Hogg Award and Lecture

Recognizing Frances H. Arnold, Ph.D.

MD Anderson Cancer Center extends congratulations to Frances H. Arnold, Ph.D., on being named the 2021 Mike Hogg Award recipient.

Arnold, who is the Linus Pauling Professor of chemical engineering, bioengineering and biochemistry at California Institute of Technology and co-chair of President Joe Biden's Council of Advisors on Science and Technology, is being recognized for pioneering directed enzyme evolution for which she was awarded the Nobel Prize in Chemistry in 2018. She has applied directed protein evolution in the areas of medicine, alternative energy and chemicals.

Arnold's lecture titled Innovation by Evolution: Bringing New Chemistry to Life will be delivered virtually this year.

The event is sponsored by the Mike Hogg Fund and hosted by MD Anderson's Division of Education and Training.

THE UNIVERSITY OF TEXAS

MD Anderson
~~Cancer~~ Center

Making Cancer History®

The dhole (*Cuon alpinus*) population has been decimated by human activities.

Edited by **Jennifer Sills**

China's dhole population at risk of extinction

The dhole (*Cuon alpinus*), a mountain canid species that once lived throughout north, south, and southeast Asia (*1*), is endangered worldwide (*2, 3*). Now found primarily in the forests, shrub belts, and grasslands of southeastern China (*4*), the dhole population includes only 949 to 2215 individuals with reproductive capacity (*3*). Although there have been positive steps toward the species' conservation, more must be done to prevent its extinction.

There are many reasons for the catastrophic decrease in the number of dholes (*5*). In the past few decades, human production and activities have occupied large areas of the dhole's original habitats (*6*). As humans have expanded into their habitats, dholes—which form groups to hunt large herbivores—have also preyed on livestock, causing conflicts (*6*). In response, humans have begun to hunt dholes (*6*), further decreasing the species' population numbers.

The dhole, which by preying on herbivores indirectly limits consumption of plants, is important to maintaining the structural and functional integrity of the ecosystem (*7*). However, there is very little protection of, or scientific research on, the species (*8*). On 5 February, the Chinese government upgraded dholes from Class II to Class I, the highest level of protected species, on the National Key Protected Species List (*9*). This classification change, although a good first step, is insufficient. We must also raise awareness about the importance of protecting dholes to discourage humans from destroying their habitats and killing them during conflicts or hunting. Governments, including China's (*10*), must immediately implement the strictest possible wildlife control measures. In addition, studies on the species' feeding habits, population dynamics, habitat selection, reproductive ecology, and genetic diversity should be conducted. Once such research yields results, we will be able to implement targeted conservation measures both in the wild and in captivity.

Xiaoyang Wu[1], Qinguo Wei[1], Sai Deni[2], Honghai Zhang[1]*

[1]College of Life Sciences, Qufu Normal University, Qufu, China. [2]Butt Field Studio, Golmud, China.
*Corresponding author.
Email: zhanghonghai67@126.com

REFERENCES AND NOTES

1. X. Wu *et al.*, *Appl. Microbiol. Biotechnol.* **100**, 3577 (2016).
2. L. S. Durbin *et al.*, in *Canids: Foxes, Wolves, Jackals and Dogs*, C. Sillero-Zubiri, M. Hoffmann, D. W. Macdonald, Eds. (International Union for Conservation of Nature, 2004).
3. J. F. Kamler *et al.*, *Cuon alpinus* (IUCN Red List of Threatened Species, 2015).
4. A. B. Venkataraman, R. Arumugam, R. Sukumar, *J. Zool.* **237**, 543 (1995).
5. T. Ma, Y. Hu, M. Wang, L. Yu, F. Wei, *Natl. Sci. Rev.* **8**, nwaa265 (2020).
6. A. Srivathsa *et al.*, *Sci. Rep.* **9**, 3081 (2019).
7. P. D. Moehlman, in *Carnivore Behavior, Ecology, and Evolution*, J. L. Gittleman, Ed. (Springer, Boston, MA, 1989), pp. 143–163.
8. X. Wu *et al.*, *Front. Microbiol.* **11**, 575731 (2020).
9. The Department of Forestry and Grassland, "The National Key Protected Species List" (2021); www.forestry.gov.cn/main/5461/20210205/122418860831352.html [in Chinese].
10. The Department of Forestry and Grassland, "The strictest wildlife control measures will be implemented" (2020); www.forestry.gov.cn/main/72/20200228/085955671954206.html [in Chinese].

10.1126/science.abi8889

Hopes for Russia's new forest code

Russia is home to one-fifth of the world's forest lands, about 90% of which comprise fragile boreal areas (*1, 2*). These forests sustain unique human cultures, valuable wilderness, and biodiversity (*1, 2*). They also play a crucial role in the flows of atmospheric moisture on which millions of people depend and sequester a substantial, though uncertain, quantity of greenhouse gases (*1*). Despite their importance, Russian forests have suffered from wasteful over-harvesting of accessible timber, inadequate protection, fire, and pests, and new challenges—such as forests on melting permafrost—are emerging (*1–3*). The previous forest code and a long history of extraction without investment have been widely blamed for exacerbating these problems (*1, 4, 5*). The Forest Council of the Russian Academy of Sciences has initiated work on a new national forest code (*2, 6*). The future of Russia's vast forests depends

on the implementation of effective policy.

Initial drafts of the new code offer a step forward. They establish sustainable policies that are based on science, institutional reforms, and good governance, and they give conservation and climate concerns nearly as much emphasis as timber revenues (*1*, *2*, *6*). However, to achieve its goals, the finalized code must be passed as law and attract ample financial and political support.

The international community can help Russia develop and support an effective forest code and the institutions required to protect its forest lands. Globally, there is valuable experience and expertise that can guide best practice. This will be essential to guide the transition to a market economy, with payment for ecosystem services and the effective implementation of conservation and sustainable management in a fast-changing world (*6*). For example, Russians with local experience could benefit from the international community's knowledge of conservation planning and effectiveness (*7–10*), the financial mechanisms to fund and incentivize forest conservation and restoration (*11*), and the development and implementation of the climate-smart strategies and practices required to sustain forests and forestry under climate change (*12*).

We urge the international forest science community to engage constructively with Russia as the country develops and applies new forest policies. Broad collaborations can help Russia develop and implement forest protection and sustainable management through an open exchange of ideas among local and international institutions. Russia's forests are the country's sovereign concern, but the fate of those forests has global consequences. International attention, collaboration, and support can help ensure that they continue to thrive.

Douglas Sheil[1,2]*, Gert-Jan Nabuurs[1], Anatoly Shvidenko[3,4]
[1]Wageningen University and Research, Wageningen, Netherlands. [2]Center for International Forestry Research, Bogor, Indonesia. [3]International Institute for Applied System Analysis, Laxenburg, Austria. [4]Institute of Forests, Siberian Branch of the Russian Academy of Sciences, Krasnoyarsk, Russia.
*Corresponding author. Email: douglas.sheil@wur.nl

REFERENCES AND NOTES
1. P. Leskinen *et al.*, Eds., "Russian forests and climate change: What science can tell us" (European Forest Institute, Finland, 2020).
2. Food and Agriculture Organization of the United Nations (FAO), "The Russian Federation forest sector: Outlook study to 2030," A. Petrov, M. Lobovikov, Eds. (FAO, Rome, 2012).
3. L. A. Henry, M. Tysiachniouk, *For. Pol. Econ.* **90**, 97 (2018).
4. S. V. Dianov, M. B. Antonov, V. A. Rigin, *Econ. Soc. Changes Facts, Trends, Forecast* **13**, 106 (2020).
5. I. V. Shutov, *Devastation and Renascence of Forestry in Russia* (Saint Petersburg Forestry Research Institute, Saint-Petersburg, 2003) [in Russian].
6. Y. N. Gagarin, *For. Sci. Iss.* **3**, 1 (2020).
7. M. G. E. Mitchell *et al.*, *Environ. Res. Lett.* **16**, 014038 (2021).
8. V. M. Adams *et al.*, *Ambio* **48**, 139 (2019).
9. D. Stralberg, C. Carroll, S. E. Nielsen, *Conserv. Lett.* **13**, e12712 (2020).
10. World Wildlife Fund, "WWF and Mondi: Towards sustainable forestry in Russia's boreal forest" (2015); https://wwf.panda.org/?249032/WWF-and-Mondi-towards-sustainable-forestry-in-Russias-boreal-forest.
11. A. Angelsen *et al.*, Eds., *Transforming REDD+: Lessons and New Directions* (Center for International Forestry Research, Bogor, Indonesia, 2018).
12. G.-J. Nabuurs *et al.*, *Forests* **8**, 484 (2017).

Long-term investment in shark sanctuaries

Shark sanctuaries, defined as national prohibitions on the commercial fishing and trade of sharks, provide globally important refuges for some shark species (*1*, *2*). The Maldives has earned international acclaim for being a sanctuary and reaps considerable economic benefit from shark dive tourism (*3*, *4*). Despite its success in shark conservation, the country may repeal or weaken its sanctuary legislation based on the desire to benefit economically from shark fishing and perceptions of increased conflict between fishers and sharks due to depredation (when sharks take or damage catches) (*5*). Although recent pandemic-related disruption of tourism likely fueled this initiative, it may have been inevitable given that revenue generated from shark tourism does not directly benefit communities that once fished sharks and now experience depredation (*3*, *4*).

Several of the world's 17 shark sanctuaries face issues similar to the one in the Maldives (*6*), suggesting that internal support for sanctuaries may erode as they become successful (*7*). In sanctuaries where overall public approval is strong, people living alongside sharks need better support. Sanctuary resilience requires targeted assistance to offset real costs to former shark fishers and those experiencing depredation (*3*). The costs of such support would be far outweighed by the national benefits of sanctuaries (*3*).

Funding for these offsets can come from a variety of sources, including intergovernmental aid from high-income, pro-shark conservation nations and sanctuary fees collected from dive tourists, routed by the government to affected communities (*3*, *4*). Likewise, the tourism industry itself, especially facility owners, could engage with these communities to include them as sanctuary beneficiaries. Nondestructive approaches for reducing depredation need to be investigated and disseminated. Funders, governments, nongovernmental organizations, the tourism sector, and conservation advocates should view shark sanctuaries not solely as policy victories but also as long-term investments with high conservation potential (*2*). These sanctuaries depend on sustained community benefits and engagement after the policy is passed and the last celebratory social media announcement is posted.

Demian D. Chapman[1,2]*, Khadeeja Ali[2,3], M. Aaron MacNeil[4], Michelle R. Heupel[5], Mark Meekan[6], Euan S. Harvey[7], Colin A. Simpfendorfer[8], Michael R. Heithaus[2]
[1]Sharks and Rays Conservation Program, Mote Marine Laboratory, Sarasota, FL 34236, USA. [2]Department of Biological Sciences, Florida International University, North Miami, FL 33181, USA. [3]Maldives Marine Research Institute, 2020 Malé, Maldives. [4]Ocean Frontier Institute, Department of Biology, Dalhousie University, NS, LSC 7088, Canada. [5]Integrated Marine Observing System, University of Tasmania, Hobart, TAS 7001, Australia. [6]Australian Institute of Marine Science, Crawley, WA 6009, Australia. [7]School of Molecular and Life Sciences, Curtin University, WA 6102, Australia. [8]Centre for Sustainable Tropical Fisheries and Aquaculture & College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia.
*Corresponding author.
Email: dchapman@mote.org

REFERENCES AND NOTES
1. C. A. Ward-Paige, B. Worm, *Glob. Environ. Change* **47**, 174 (2017).
2. M. A. MacNeil *et al.*, *Nature* **583**, 801 (2020).
3. K. Ali, H. Sinan, *J. Mar. Biol. Ass. India* **56**, 34 (2014).
4. J. Zimmerhackel *et al.*, *Mar. Pol.* **100**, 21 (2019).
5. "Maldives plans to legalise shark fishing," *South Asia Monitor* (2021).
6. N. McKenzie, "BCFA recommends govt. temporarily lift shark harvesting ban," *Eyewitness News* (2020).
7. J. K. Carlson *et al.*, *Environ. Conserv.* **46**, 264 (2019).

Science for Seminaries faculty visit George Washington University's Mind-Brain Institute in 2018.
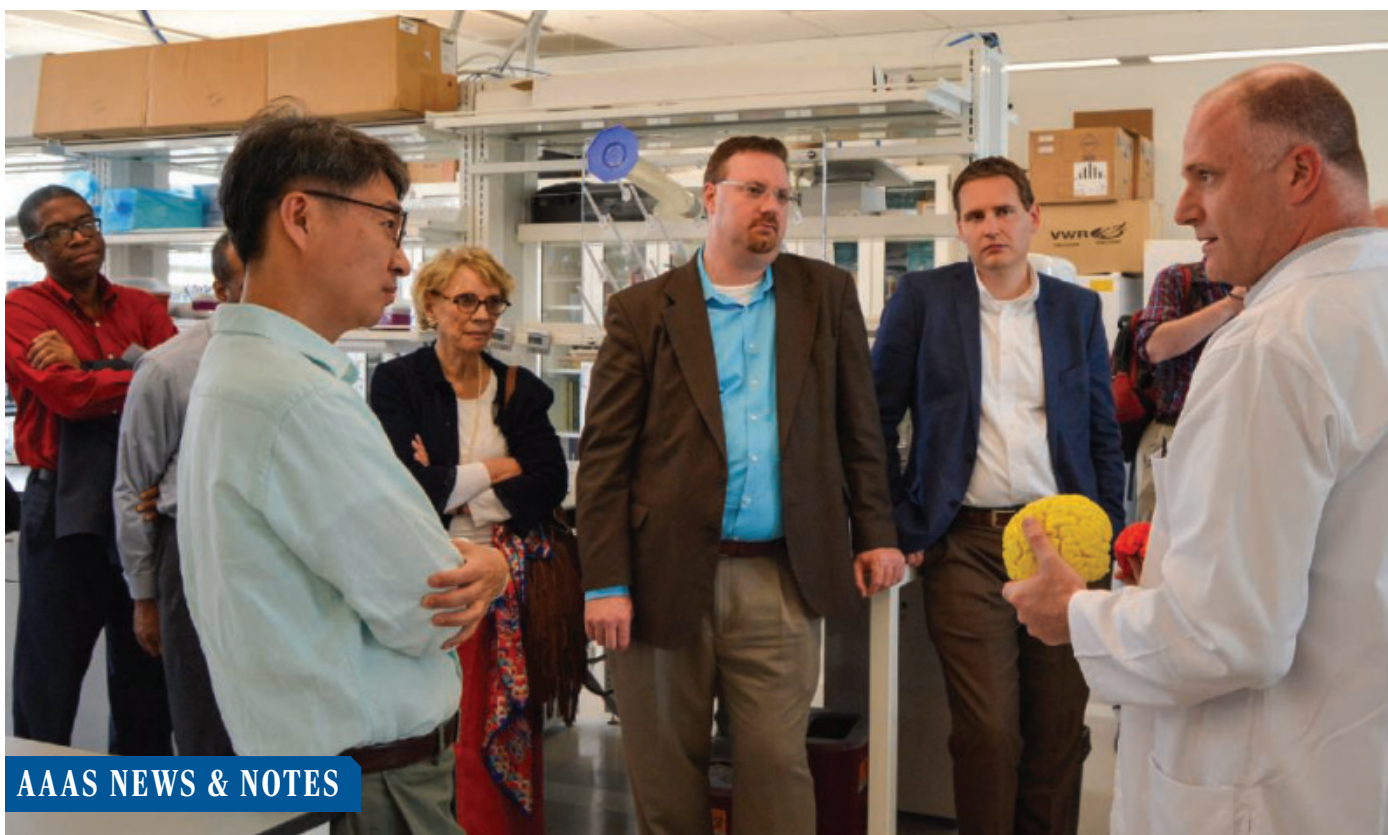
AAAS celebrates 25 years connecting science and religion

Relationship building is key for the Dialogue on Science, Ethics, and Religion

By **Becky Ham**

On 15 June, the AAAS Dialogue on Science, Ethics, and Religion (DoSER) program will celebrate its 25th anniversary with a program of speakers and group discussions, covering topics from artificial intelligence to racism. While DoSER (www.aaas.org/DoSER) has had notable successes in building relationships between religious and scientific communities, the anniversary is a prompt to look forward, not back, said the program leaders.

At the event, called "Forward Together: Where Science, Ethics, and Religion Intersect in a Changing World," speakers "will talk about the issues that are hot topics in science and technology today that have a broad impact on life around the globe," said DoSER Director Jennifer Wiseman, "and how faith communities are integral to good uses of science and technology going forward."

The program has changed over the past quarter-century in several key aspects, mostly in expanding its outreach to new communities and enriching the resources available to promote the religion-science dialogue. The topics in science, technology, and ethics that have animated this dialogue have also changed over time, from early discussions about human origins and evolution to climate change, environmental stewardship, public health, artificial intelligence, neuroscience, genetics, science and societal justice, and more.

"But even now, our emphasis is not primarily on the topics," Wiseman said. "It's on building connections between people, so that scientists and faith communities get a chance to know each other."

Part of the foundation for building these connections came from DoSER's Perceptions Project, concluded in 2015, that investigated the beliefs scientists and religious leaders hold about each other. In a series of workshops held in six U.S. cities, members of both communities met—sometimes for the first time—to build the architecture for new relationships.

There is "sometimes a knee-jerk presumption that religious communities are going to be uninterested or even a little reticent to enter discussions involving science, and that is just not the case," Wiseman said.

Science for Seminaries, a DoSER project active since 2013, helps U.S. and Canadian theological seminaries incorporate relevant science into their core curricula. The goal of the immensely popular program is to give future faith leaders the tools to engage with their congregations about science and technology issues that permeate nearly every aspect of life.

The final cohort of seminaries begins the program this spring, and DoSER hopes to continue the program with future funding. While the seminaries program has focused on Christian communities, DoSER's outreach has expanded over time to reach out to other faiths as well.

One of DoSER's partners is Sinai and Synapses, a program to bridge the Jewish faith and science led by Rabbi Geoffrey A. Mitelman. For the group's Scientists in Synagogues initiative, "DoSER's

reputation and ability to engage scientists in high-level religious conversations have been instrumental in our success," said Mitelman.

"As a layperson, it can be hard to distinguish scientific information that's accurate from what's well-meaning but inaccurate—let alone when willful deception enters the picture," he added. "Being able to turn to a trusted source has been incredibly valuable, and as we engage Jewish communities in thinking about science, DoSER has been a tremendous resource."

Scientists' interest in engaging with religious communities "has grown dramatically" over the years, said Wiseman. In response, DoSER launched initiatives such as the Engaging Scientists in the Science and Religion Dialogue project, which provides connections and resources for scientists to build their relationships with the religious public.

Nalini Nadkarni, a University of Utah forest ecologist and conservationist who participated in the Engaging Scientists project, wanted to reach out to people who might not visit museums or botanical gardens. "I learned that over 80% of people self-identify as being religious or believing in God," she said. "So I thought that if I could find common ground with faith-based groups about the positive values of trees for people, I would have a chance for finding allies to protect trees and nature."

Nadkarni now speaks to religious groups from a new text that she created by combing through the scriptures of the world's religions, sharing spiritual values that have been placed on trees and forests by these communities.

Her work with faith groups has reinforced an idea that she carries over into her outreach with other groups such as the incarcerated, urban youth, and artists, she said. "That is, a scientist must feel and show intellectual humility to be successful in engaging 'the other,'" Nadkarni said. "Sometimes we have to put aside for the moment our own precepts and assumptions enough to truly listen, truly hear what the other group is articulating."

DoSER has launched a new website, sciencereligiondialogue.org, that contains resources such as profiles of scientists who effectively engage with religious students and communities and short videos from the Science: The Wide Angle series. The series features scientists speaking about topics that also resonate with religious communities, like the "awe and wonder" inherent in the natural world, said Katy Hinman, DoSER's associate program drector.

"The point is that no matter what you believe personally as a scientist, there are ways you can engage positively with people of faith all around you that welcome them into science and don't create unnecessary barriers," Hinman said.

DoSER's role is significant in AAAS diversity, equity, and inclusion efforts, both in and out of the organization. Several DoSER team members are active with the association's staff-led Inclusion, Diversity, Equity, and Accessibility (IDEA) committee. "Engagement with religious communities can increase and support diversity in the sciences," said Program Associate Lilah Sloane-Barrett, who also cochairs the committee. "Some U.S. polls show that people in communities underrepresented in STEM fields identify as more religious than those already well-represented in the sciences," she noted.

A strength of the DoSER program is that it does not "evaluate religious belief," said Wiseman. "You don't have to believe in everything that a faith community believes or practices in order to have a positive interaction and relationship regarding science."

In symposia at the AAAS Annual Meeting, in its popular holiday lectures, and in the events that the program holds around the country, DoSER is continually searching for new ways to facilitate relationships between religious and scientific communities. But the "focus on building connections and longer-term relationships is what outlasts specific projects," said Wiseman.

RESEARCH

ISLAND ECOLOGY

Accelerating ecosystem disruption

Oceanic islands are among the most recent areas on Earth to have been colonized by humans, in many cases in just the past few thousand years. Therefore, they are important laboratories for the study of human impacts on natural vegetation and biodiversity. Nogué *et al.* provide a quantitative palaeoecological study of 27 islands around the world, focusing on pollen records of vegetation composition before and after human arrival. The authors found a consistent pattern of acceleration of vegetation turnover after human invasion, with median rates of change increasing by a factor of six. These changes occurred regardless of geographical and ecological features of the island and show how rapidly ecosystems can change and how island ecosystems are set on new trajectories. —AMS *Science*, this issue p. 488

Human arrival on islands and subsequent activities, such as pineapple farming in Hawai'i pictured here, have accelerated vegetation turnover in these isolated ecosystems.

INDUCED SEISMICITY
Varying the stop lights

Traffic light protocols can help to mitigate induced earthquakes from unconventional oil production. However, they are not geographically tuned to account for how shaking may actually translate to structural damage. Schultz *et al.* incorporated damage tolerance into a traffic light protocol for the Eagle Ford shale play. They found that shut-off may be necessary more quickly in populated regions, whereas sparsely populated areas of the play can take up to a magnitude 5 earthquake without issue. This risk-based strategy provides a more nuanced approach to regulating induced seismicity. —BG

Science, this issue p. 504

PALEOECOLOGY
Pre-Columbian reforestation in Amazonia

An early 17th-century temporary reduction in global atmospheric carbon dioxide (CO_2) levels has now been attributed to reforestation in Amazonia after the catastrophic loss of life of the indigenous population caused by diseases brought by European invaders. Using fossil pollen data from Amazonian lake sediments with temporal resolution over the past millennium, Bush *et al.* found that forest recovery began 300 to 600 years before the population crash. The more recent nadir in atmospheric CO_2 was not associated with rapid reforestation at that time. The vegetation changes appear to be the result of changing patterns of land use in the centuries preceding the European arrival and the resulting devastation, whereas the cause of the CO_2 decline remains enigmatic. —AMS

Science, this issue p. 484

MULTIFERROICS
An optically active spiral

The material cupric oxide exhibits magnetoelectric coupling, meaning that its magnetic properties can be controlled by electric fields. In its spin spiral phase, cupric oxide has a spiral magnetic ordering that can be right- or left-handed. Masuda *et al.* used electric fields to create purely left- or right-handed samples and then studied their optical activity. The samples exhibited natural optical activity, which the researchers were then able to control with electric fields. —JS

Science, this issue p. 496

NANOFLUIDICS
Gated ion flow in graphene oxide membranes

Cells are adept at fast, gated ion flow through tailored channels, which is key to many biological processes. Xue *et al.* developed ion transistors from reduced graphene oxide membranes and observed a field-enhanced diffusivity of the ions (see the Perspective by Hinds). By applying electrical gating, the average surface potential on the graphene layer could be controlled, thus altering the energy barrier for ion intercalation into the channel and leading to very high diffusion rates. The authors observed selective ion transport

two orders of magnitude faster than the ion diffusion in bulk water. —MSL

Science, this issue p. 501; see also p. 459

ACUTE LUNG INJURY
ROS-mediated lung protection

Neutrophils accumulate in the lung after acute lung injury (ALI) and play a role in the innate immune response through multiple mechanisms, including the production of reactive oxygen species (ROS). Yuan *et al.* investigated the mechanisms regulating ROS production during ALI and developed a therapeutic intervention. The authors found that the kinases MAP3K2 and MAP3K3 inhibited ROS production from neutrophils. Pazopanib, a specific MAP3K2/3 inhibitor, ameliorated ALI in mice by modulating phosphorylation of p47, a subunit of the NADPH oxidase 2 complex. This treatment was effective in reducing pulmonary edema in a pilot study in patients who underwent lung transplantation. —MM

Sci. Transl. Med. **13**, eabc2499 (2021).

ECOSYSTEM ENGINEERS
Digging for water

Water is scarce in dryland ecosystems. Some larger animals in these regions dig wells that may provide water to other species. This behavior may have been common among megafauna that are now extinct, especially in North and South America, where megafaunal extinctions were the most severe. Lundgren *et al.* tested whether feral equids (horses and donkeys) reintroduced to desert regions in the North American southwest dig wells that provide ecosystem-level benefits. They found that equid-dug wells increased water availability, were used by a large number of species, and decreased distance between water sources. Abandoned wells also led to increased germination in key riparian tree species. Such equid-dug wells improve water availability, perhaps replacing a lost megafaunal function. —SNV

Science, this issue p. 491

IMMUNOLOGY
Putting DCs into overdrive

When dendritic cells (DCs) detect signals indicative of infection, cell death, or cancer, they respond by activating a signaling complex known as the inflammasome, which results in proinflammatory cytokine secretion but usually leads to the death of the DCs. By analyzing human primary DCs from the blood, spleen, and bone marrow, Hatscher *et al.* found that type 2 conventional DCs, in an unusual turn, did not die after inflammasome-induced cytokine secretion. Instead, these cells entered a hyperactive state that elicited more effective responses from certain T helper cell subsets. —LKF

Sci. Signal. **14**, eabe1757 (2021).

Horses and donkeys dig shallow wells that are used by many species.

IN OTHER JOURNALS

Escherichia coli uses several defense systems against T4 phage parasites, one of which is shown here in a colored electron micrograph image.

PHAGE DEFENSE
toxIN death knell

Bacteria have diverse mechanisms to protect themselves against phage. Some, such as the CRISPR-Cas system, are always ready to recognize and eliminate invaders. Others, such as toxin-antitoxin systems, are only activated after phage infection. Guegler and Laub investigated the *toxIN* system that protects *Escherichia coli* against several bacteriophages. The toxin, toxN, is a ribonuclease and the antitoxin, *toxI*, is an RNA with an array of repeats. Under normal conditions, toxN cuts *toxI* and binds the single motif. Infection by the phage T4 shuts off host transcription, including transcription of *toxIN*. Because *toxI* is unstable, toxN is released to cleave mRNA in the cell, which by this time is mainly phage derived. This prevents the production of new phage particles. Thus, by appropriating the host's replicative machinery, phage also risk releasing their *toxIN* nemesis. —VV

Mol. Cell 10.1016/j.molcel.2021.03.027 (2021).

3D GENOME
"Slinky" chromatin in archaea

Only eukaryotes and archaea use histones to package their DNA. This observation has prompted suggestions of an evolutionary affinity between these two domains of life. However, there are many differences between the structure of histones between the domains. Bowerman *et al.* extended earlier work to show how archaeal histones store and unpack DNA. In eukaryotes, a packet of four pairs of histones wraps around every ~147 base

Graphical representation of a nanographene fiber with a pitched double helix

NANOCARBON SCIENCE

One-dimensional self-assembly of curved NGs

Since the first experimental detection of graphene, a huge research effort has been made to implement its exceptional properties in carbon-based technologies. The well-ordered molecular alignments using small pieces of graphene called nanographenes (NGs) can be used to engineer supramolecular structures for certain practical applications. Kato *et al.* report the discovery of a negatively curved NG that is able to spontaneously self-assemble in one dimension by continuous π–π stacking without any assisting substituents. This NG forms double-helix supramolecular all-sp^2-carbon nanofibers that demonstrate high gelation efficiency. This work paves the way for the design and synthesis of supramolecular materials based on analogous NGs that have great potential in future applications. —YS *J. Am. Chem. Soc.* **143**, 5465 (2021).

pairs of DNA in a structure called a nucleosome. By contrast, the archaeal equivalent of a nucleosome, the archaeasome, forms a histone core with more than four histone pairs. Archaeasomes can expand, in effect stretching the coil, to open up the DNA in a way that is very different from how this process occurs in eukaryotes. —DJ

eLife **10**, e65587 (2021).

IMMUNOMETABOLISM

Mobilizing metabolism against HIV

HIV-1 alters the metabolism of infected immune cells, targeting lipid, tryptophan, and glucose metabolic pathways by mechanisms that remain poorly understood. Guo *et al.* used transcriptomics and proteomics to discern the details. HIV-1 infection enhanced associations between NLRX1, a

mitochondrial antiviral pattern recognition receptor, and FASTKD5, a protein that modulates mitochondrial energy use. This change increased oxidative phosphorylation (OXPHOS) and glycolysis, which in turn promoted viral replication. HIV-1 replication in vitro became compromised when OXPHOS was inhibited by the antidiabetic drug metformin. Thus, supplementation of combination antiretroviral therapy with metformin and other drugs targeting these metabolic pathways could be helpful in the treatment of HIV-1. —STS

Nat. Immunol. **22**, 423 (2021).

EVOLUTION

Relative immunity

Humans and other apes tend to respond differently to infection compared with other primates. For instance, baboons can tolerate almost 10-fold greater

exposure to bacterial cell wall material than humans can before developing septic shock. Hawash *et al.* suspected that innate immune signaling lies at the heart of the differences observed among different types of primates. The authors looked at whole-genome expression patterns of leukocytes from four diverse primates within 24 hours of responding to viral and bacterial pathogens. Indeed, the apes' innate responses were first off the mark. This finding indicates that apes undergo immediate pathogen clearance despite the risk of potential collateral damage to self. This is possibly because, evolutionarily, this strategy favors the longer-lived apes, who experience more pathogen exposure during their lifetimes. —CA

Proc. Natl. Acad. Sci. U.S.A. **118**, e2015855118 (2021).

CONFINED CATALYSIS

Carbon nitride nanotube reactors

Confinement of catalyzed reactions in nanoscale spaces can lead to higher activity and selectivity and may also help to stabilize catalysts. Zou *et al.* fabricated a free-standing membrane of carbon-nitride nanotubes (~40-nanometer pore diameter) by polymerizing melamine in anodized aluminum membrane templates. After etching and freeing the membrane, exposure to visible light generated electron-hole pairs that could be used to degrade methylene blue or, with the addition of gold nanoparticles in the pores, to oxidize amines such as benzylamine in imines in the presence of oxygen. The higher reactivity relative to bulk catalysts was attributed to a strong electric field effect on substrates, longer photogenerated charge lifetimes, and enhanced fluid flow and oxygen diffusion within the channels. —PDS

ACS Nano 10.1021/acsnano.0c09661 (2021).

PSYCHOLOGY

Structural whitening

As the US population becomes more racially diverse, it is unclear how ethnic white populations will respond to these demographic changes. Anicich *et al.* found experimentally that when white Americans were given the opportunity to populate fictional cities, they imposed greater racial segregation in areas that they frequented more often, such as work or school, because they feel greater anxiety around non-whites. In a follow-up study, the authors examined policies at tennis and golf clubs across the United States, and found that in areas with higher racial diversity, clubs engaged in more exclusionary behavior, such as enacting strict dress codes. These findings suggest that as racial diversity increases, white Americans may respond by trying to structure their environment in more segregated ways. —TSR

*J. Exp. Soc. Psychol.* **95**, 104117 (2021).

REVIEW SUMMARY

PLANT SCIENCE

Chronoculture, harnessing the circadian clock to improve crop yield and sustainability

Gareth Steed†, Dora Cano Ramirez†, Matthew A. Hannah, Alex A. R. Webb*

BACKGROUND: Global population growth, increased meat consumption, and concerns about sustainability of resources are driving forces for future agriculture to produce higher yields with reduced inputs. Studies, mostly performed in the model plant *Arabidopsis thaliana*, demonstrate that the circadian clock, which is a 24-hour oscillator that is an adaptation to living on a rotating planet, has profound effects on the physiology and development of plants. The circadian system contributes to the regulation of flowering, biomass, photosynthesis, water use, temperature stress responses, and pathogen defenses, which are important yield components in plant crops.

ADVANCES: Circadian timing in plants is controlled by circadian oscillators in every cell. These oscillators contain pathways of regulation with high degrees of feedback between transcriptional regulators that are expressed in a temporal series through the day and night cycle. The genetic architecture of these circadian oscillators has been largely solved in plants, animals, and fungi during the last quarter-century. This has revealed a plant-specific set of circadian oscillator genes that are common between the major crops. Genome sequencing of crop plants and genetic mapping studies have demonstrated that agricultural breeding has selected for allelic variants in circadian genes, particularly for agricultural modification of the time to flowering. At the same time, indoor crop growth in controlled environments has expanded opportunities to improve agriculture by simultaneously engineering the external environmental cycles and the internal circadian cycles of crops. This knowledge of the genetic structure of the circadian system; the discoveries that the plant circadian oscillators regulate a swathe of plant physiology, metabolism, and gene expression; and improved genetic tools mean that it is now possible to consider translational research that aims to use the circadian system as an approach to improve crop yield and reduce inputs.
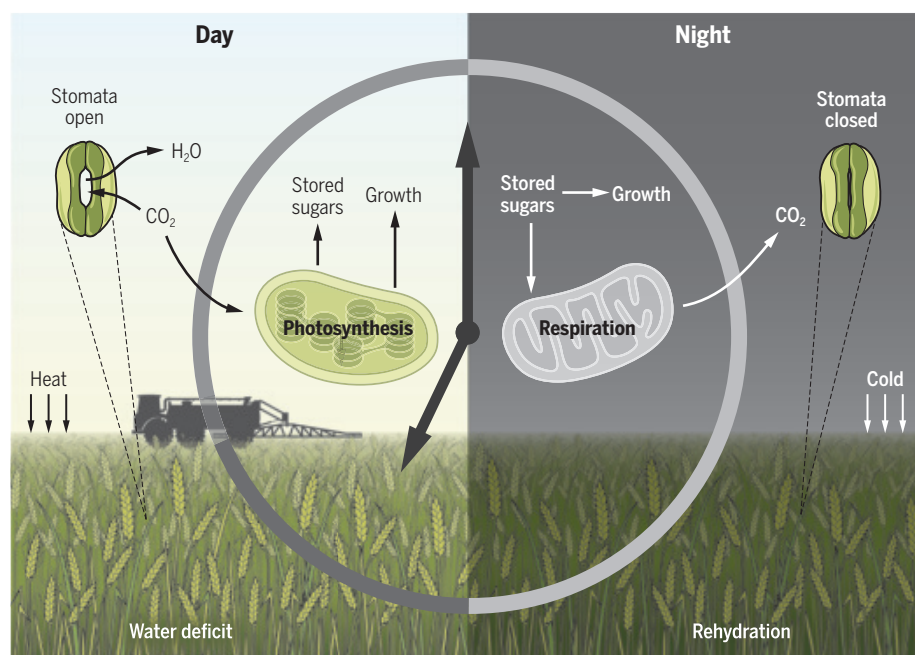
OUTLOOK: Circadian biology might have translational impact in crop science through the practice of "chronoculture," in which the daily timing of agronomic interventions and the genetic basis of circadian rhythms are targets for crop improvements and reduced inputs. In the short term, this would include the use of automation for round-the-clock monitoring and control of plant crop growth and health. Drones and low-cost single-board computers can capture high-frequency data about rhythms of photosynthesis, gas exchange, water use, stomatal movements, and growth of crops in the field to inform mechanistic and agronomic models of crop growth. Experiments in model systems have demonstrated daily rhythms in the responsiveness of plants to temperature, water, pests, and herbicides. With additional data from studies in crop species on the farm, it might be possible to develop new agronomic approaches in which interventions with nutrients, chemicals, light, and temperature are applied at the time when they are most effective. Breeding new germplasm with greater variation at circadian alleles might contribute to improved growth and sustainability. In the future, gene editing and transgenic approaches could be used to engineer crops. For example, plants with slow-running circadian clocks might be better adapted to high-latitude summers. New tissue-specific gene expression strategies that allow temporal and spatial control of transgene expression represent long-term targets for engineering crops with optimized traits. For example, alteration of circadian function specifically in phloem companion cells could be used to engineer new flowering-time traits without affecting other circadian outputs. By using split protein approaches, it should be possible to express proteins at specific times of day in desired tissues to confer stress tolerance or produce bioproducts for harvest, with less negative effects than constitutive expression. Breeding and transgene approaches are expensive because they are long-term programs subject to regulatory scrutiny. Consequently, there is a need for increased fundamental research and applied agronomic studies investigating the basic mechanism and applied uses of circadian biology in agriculture. Major goals of chronoculture research will be measuring, in the major crops in the field, the impact of the circadian oscillator on yield-relevant traits and quantifying the effect of time of day on agronomic treatments. ■

Wheat growing in day and night cycles of light and dark and warm and cold. In the light (left), plants photosynthesize, fixing carbon for growth and the production of transient reserves of sugars. At night (right), plants consume stored sugars. Circadian clocks anticipate these environmental cycles and regulate most plant biology.

CREDIT: KELLIE HOLOSKI/*SCIENCE*

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abc9141

RESEARCH ARTICLE SUMMARY

TRANSCRIPTION

Structural insights into preinitiation complex assembly on core promoters

Xizi Chen†, Yilun Qi†, Zihan Wu†, Xinxin Wang†, Jiabei Li†, Dan Zhao†, Haifeng Hou†, Yan Li, Zishuo Yu, Weida Liu, Mo Wang, Yulei Ren, Ze Li, Huirong Yang, Yanhui Xu*

INTRODUCTION: RNA polymerase II (Pol II)–mediated transcription initiation requires assembly of a preinitiation complex (PIC), during which the 14-subunit transcription factor IID (TFIID) recognizes core promoters and recruits TFIIA, TFIIB, TFIIE, TFIIF, TFIIH, and Pol II to sequentially assemble core PIC (cPIC), intermediate PIC (mPIC), and holo PIC (hPIC). The textbook model of transcription initiation starts with binding of TBP (TATA box–binding protein, a TFIID subunit) to TATA box. However, up to 85% of coding genes lack consensus TATA box, and the TFIID complex is required for almost all Pol II–mediated gene transcription. Despite extensive structural studies of TBP-based PIC on TATA box promoters, it remains elusive how TFIID supports PIC assembly on different promoters.

RATIONALE: We reconstituted and determined cryo–electron microcopy structures of the human TFIID-based PIC. Stepwise PIC assembly was characterized by 25 complex structures in different compositional and conformational states on 13 different (natural, composite, and mutant) promoters. Structures of Pol II, TFIID modules, and TBP-promoter at near-atomic resolution permit detailed structural analyses.

RESULTS: Structures in distinct conformations reveal a shared TFIID-binding pattern and loading of TBP to TATA and TATA-less promoters. Unexpectedly, TBP similarly bends TATA box and TATA-less promoters in PIC. PIC assembly on different promoters diverges into two tracks at cPIC and converges at hPIC. On track I, cPIC, mPIC, and hPIC on TATA-DBE promoters adopt the Park, Neutral, and Drive conformations, respectively, indicating stepwise promoter deposition to Pol II accompanied by extensive modular reorganization. On track II, cPIC, mPIC, and hPIC on TATA-only and TATA-less promoters adopt the Drive conformation, indicating direct promoter deposition. The differences result from distinct promoter compositions, which lead to "matched" versus "repositioned" modular separation on promoters and result in distinct PIC architectures and promoter trajectories. In hPIC, TFIID stabilizes PIC organization and supports loading of cyclin-dependent kinase 7 (CDK7) onto Pol II and CDK7-mediated C-terminal domain (CTD) phosphorylation.

CONCLUSION: Our study resolves the long-standing controversy between the lack of TATA box in most core promoters and the necessity of the TFIID complex (but not TBP alone) in transcription. TFIID recognizes promoters and supports TBP-induced promoter bending and two-track PIC assembly on highly diversified core promoters. The stepwise promoter deposition may serve as a checkpoint to prevent promiscuous initiation before PIC is fully assembled, and the hPIC offers a shared starting point for transcription initiation independent of promoter type. Structural visualization of PIC assembly provides a framework for further studies of transcription initiation in the context of transcription factors, coactivators, and epigenetic regulators. ∎
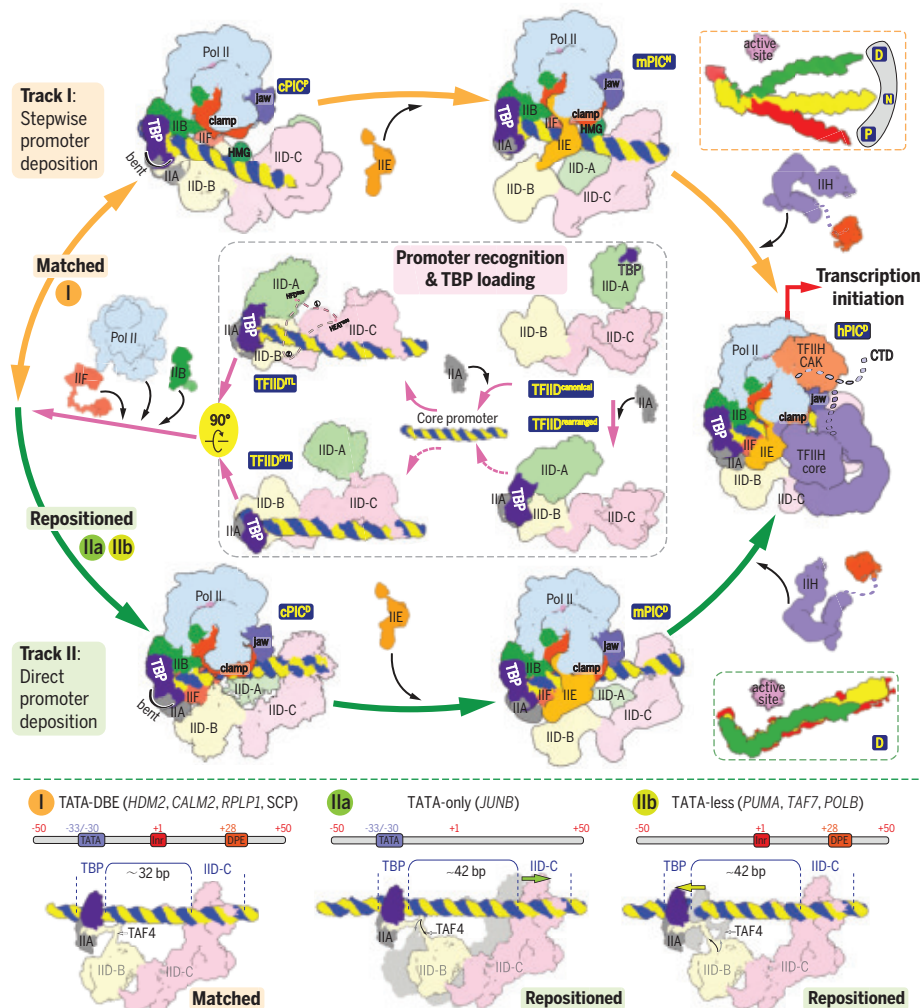
Schematic model of PIC assembly. Proposed working model of TFIID-supported promoter recognition (inner section) and two-track PIC assembly on different promoters (outer section). P, N, and D denote the Park, Neutral, and Drive conformations, respectively. Right panels (with dashed outlines): Comparison of promoter conformations in cPIC (red), mPIC (yellow), and hPIC (green). Lower panel: The matched and repositioned modular separation on promoters during cPIC assembly. Structures are derived from this study, and representative core promoters are indicated. HMG, high-mobility group; DBE, TFIID-binding element; CAK, CDK-activating kinase.

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aba8490

CRISPR

Toxin-antitoxin RNA pairs safeguard CRISPR-Cas systems

Ming Li†*, Luyao Gong†, Feiyue Cheng†, Haiying Yu, Dahe Zhao, Rui Wang, Tian Wang, Shengjie Zhang, Jian Zhou, Sergey A. Shmakov, Eugene V. Koonin, Hua Xiang*

INTRODUCTION: CRISPR-Cas systems efficiently protect bacteria and archaea from viruses and other types of foreign DNA, but, characteristically of defense systems, they also impart non-negligible fitness costs on the host, for example, the risk of autoimmunity and the repulsion to exogenous beneficial genes. Presumably, these costs result in frequent loss of CRISPR-Cas in bacteria, which is reflected in its patchy distribution, even among closely related bacterial strains. Nevertheless, in the current genome sequence databases, ~40% of bacterial and ~90% of archaeal genomes carry CRISPR-*cas* loci, suggesting the possibility that in addition to the direct benefits of adaptive immunity, mechanisms might exist that mitigate the costs of CRISPR systems and prevent their loss.

RATIONALE: We specifically looked into an archaeal type I-B CRISPR-Cas, where the genes encoding the subunits of the CRISPR effector Cascade cannot be deleted individually but can be readily deleted as a whole, including a 311–base pair intergenic region. These observations suggest that the Cascade gene cassette (*cas6-cas8-cas7-cas5*) includes a toxic component that makes it addictive to the host (elicits cellular toxicity once any of the *cascade* genes is deleted). We cloned and extensively analyzed the intergenic region between *cas6* and *cas8*, which allowed us to identify the Cascade-repressed toxin gene *creT*, along with an associating CRISPR repeat–like sequence that appears to be required for transcriptional repression of *creT*. We hypothesized that the repeat-like sequence is part of a CRISPR RNA–resembling antitoxin (CreA) RNA, which represses the toxin jointly with Cascade. We reasoned that CreTA would make the *cascade* genes addictive for the host.

RESULTS: The intergenic sequence between *cas6* and *cas8* caused toxicity in cells lacking one or more *cascade* genes. By extensive mutational analysis, we identified the RNA toxin gene *creT* and its critical elements, namely, a combination of a strong Shine-Dalgarno motif, an efficient start codon, two minor arginine codons (AGA) located immediately downstream, and a stable stem-loop structure. Overexpression of tRNA^{UCU} relieved the toxicity of CreT, supporting a mechanism whereby this RNA toxin arrests cellular growth by sequestering the rare arginine tRNA^{UCU}.

Mutational analysis of *creT* and its neighboring sequences revealed an adjacent CRISPR repeat–like sequence that is required to suppress the toxicity of CreT. This repeat-like sequence is immediately followed by a spacer-like sequence and a transcription terminator. By Northern blotting and RNA sequencing, we validated the expression of CreA RNA, a CRISPR RNA variant that lacks a 3′ handle. The spacer of CreA partially matches the promoter of *creT* (P$_{creT}$), and using a reporter gene, we confirmed that CreA, as a complex with Cascade, represses P$_{creT}$. Similar to CRISPR interference, repression of *creT* requires a protospacer adjacent motif (PAM) and the PAM-proximal base pairing. In cells lacking CreTA, the *cascade* genes become susceptible to disruption by transposable elements. Our bioinformatic analysis identified several CreTA analogs associated with diverse archaeal and bacterial CRISPR-*cas* loci and containing PAMs corresponding to those of the respective CRISPR systems. Notably, these CreTA analogs hold little conservation in nucleic acid sequence, suggesting that they have highly divergently evolved and, conceivably, exploited different toxicity mechanisms.

CONCLUSION: Our data unearth previously unnoticed toxin-antitoxin RNA pairs that prevent the loss of CRISPR-*cas* loci by making them addictive to the host cell. The naturally occurring reprogramming of CRISPR effectors for gene regulation highlights the multifunctionality of CRISPR-Cas in bacteria and archaea and illuminates the emerging topic of the evolution of antiviral defense and gene regulation. ∎
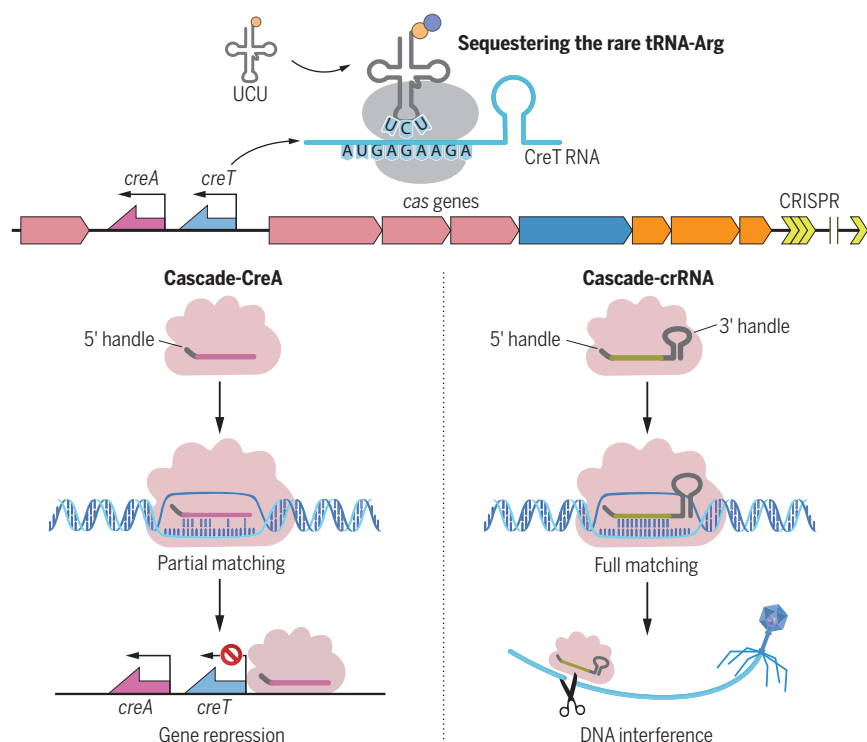
Toxin-antitoxin RNA pair CreTA safeguards CRISPR-Cas. CRISPR effector (Cascade) is not only guided by CRISPR RNA to inactivate full-matching foreign nucleic acids but is also co-opted by CreA RNA to transcriptionally repress the toxin gene *creT* through partial complementarity between CreA and the *creT* promoter. When Cascade is inactivated, the derepressed CreT RNA sequesters the rare tRNA^{UCU} that decodes a minor arginine codon and arrests cellular growth, thus making the CRISPR effector addictive to the host cell.

S **READ THE FULL ARTICLE AT**
https://doi.org/10.1126/science.abe5601

VACCINES

Modulation of MHC-E transport by viral decoy ligands is required for RhCMV/SIV vaccine efficacy

Marieke C. Verweij†, Scott G. Hansen†, Ravi Iyer†, Nessy John†, Daniel Malouli, David Morrow, Isabel Scholz, Jennie Womack, Shaheed Abdulhaqq, Roxanne M. Gilbride, Colette M. Hughes, Abigail B. Ventura, Julia C. Ford, Andrea N. Selseth, Kelli Oswald, Rebecca Shoemaker, Brian Berkemeier, William J. Bosche, Michael Hull, Jason Shao, Jonah B. Sacha, Michael K. Axthelm, Paul T. Edlefsen, Jeffrey D. Lifson, Louis J. Picker*, Klaus Früh*

INTRODUCTION: Strain 68-1 rhesus cytomegalovirus (RhCMV) vaccine vectors expressing simian immunodeficiency virus (SIV) antigens elicit immune responses that stringently control and ultimately clear highly pathogenic SIV challenge in more than half of rhesus monkeys (RMs). The high frequency of effector-differentiated CD8+ T cells elicited by this vaccine contribute to this efficacy, mediating SIV replication arrest before the establishment of long-lived SIV reservoirs. However, we unexpectedly found that SIV peptides targeted by these CD8+ T cells are presented by major histocompatibility complex-E (MHC-E) or MHC-II instead of MHC-Ia. This raised the question of how these unconventional T cell responses arise and whether they are necessary for vaccine efficacy.

RATIONALE: MHC-E binds the conserved VL9 peptide embedded in the leader sequence of MHC-Ia proteins. This MHC-E–VL9 complex predominantly engages inhibitory receptors on natural killer (NK) cells, thus serving as a protective "self" signal for healthy cells. To evade NK cells, both rhesus and human CMVs encode viral glycoproteins (Rh67 and UL40, respectively) that contain a VL9 sequence providing the MHC-E/VL9 complex to NK cells when cellular VL9 is not available because of viral MHC-Ia downregulation. We hypothesized that viral-encoded VL9 also controls the ability of 68-1 RhCMV to elicit MHC-E–restricted CD8+ T cells. This allowed us to examine the role of these cells in RhCMV/SIV vaccine efficacy.

RESULTS: We show that Rh67-embedded VL9 is required for intracellular transport of MHC-E in RhCMV-infected fibroblasts and for their recognition by MHC-E–targeting CD8+ T cells. Rh67 deletion did not substantially affect the character of RhCMV/SIV–elicited T cell responses, but such deletion, or VL9 mutation within Rh67, completely abrogated MHC-E–restricted CD8+ T cell priming, leaving vaccine-elicited T cell responses entirely MHC-II restricted. In contrast to RMs vaccinated with the parent 68-1 RhCMV/SIV vaccine, post-SIV challenge replication arrest was not observed among RMs vaccinated with the Rh67-deleted vaccine. Thus, 68-1 RhCMV/SIV vaccine efficacy requires Rh67-induced, MHC-E–restricted CD8+ T cell priming, implicating this response in the mechanism of protection.

CONCLUSION: RhCMV appears to elicit MHC-E–restricted T cells as a consequence of viral NK cell evasion. However, as separately reported, these responses are only manifest in 68-1 RhCMV because of a genetic rearrangement that abrogates the function of eight independent viral inhibitors of these responses. Thus, RhCMV has evolved to evade both NK cells and the MHC-E–restricted CD8+ T cells that result from this evasion. By contrast, SIV neither elicits nor seems to evade these responses. Our results strongly suggest that CD8+ T cells targeting MHC-E–presented HIV peptides will be necessary for an effective CMV-based HIV vaccine Fortunately, both the positive and negative RhCMV regulators of MHC-E–restricted responses are conserved in human CMV, suggesting that human CMV and/or HIV vaccines might also be programmed to elicit this unusual type of protective immunity. ∎

S READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abe9233
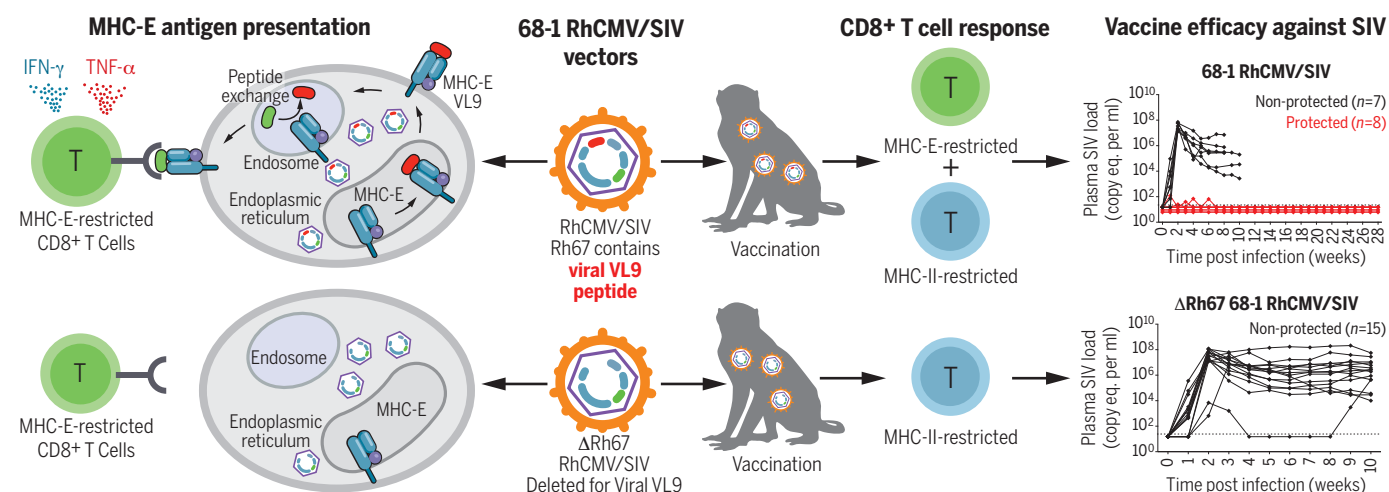
68-1 RhCMV/SIV vectors encode an MHC-E peptide ligand that enables antigen presentation by MHC-E and the induction of the MHC-E–restricted CD8+ T cells required for protection against SIV challenge. Fibroblasts infected in vitro with 68-1 RhCMV can stimulate MHC-E–restricted CD8+ T cells only if the MHC-E VL9 peptide ligand (in red) embedded within the Rh67 gene is expressed. This peptide is loaded onto MHC-E in the endoplasmic reticulum, thus enabling the intracellular transport of MHC-E to the cell surface. MHC-E is rapidly internalized from the cell surface and is followed by a (still hypothetical) exchange of VL9 with virus-derived antigenic peptides (in green). The T cell response elicited by 68-1 RhCMV vectors lacking viral VL9 is entirely MHC-II restricted and fails to protect against SIV. Thus, MHC-E–restricted CD8+ T cells are essential for 68-1 RhCMV/SIV vaccine–mediated SIV replication arrest.

PATHOGEN EVOLUTION

Stepwise pathogenic evolution of *Mycobacterium abscessus*

Josephine M. Bryant, Karen P. Brown, Sophie Burbaud, Isobel Everall, Juan M. Belardinelli, Daniela Rodriguez-Rincon, Dorothy M. Grogono, Chelsea M. Peterson, Deepshikha Verma, Ieuan E. Evans, Christopher Ruis, Aaron Weimann, Divya Arora, Sony Malhotra, Bridget Bannerman, Charlotte Passemar, Kerra Templeton, Gordon MacGregor, Kasim Jiwa, Andrew J. Fisher, Tom L. Blundell, Diane J. Ordway, Mary Jackson, Julian Parkhill*, R. Andres Floto*

INTRODUCTION: Nearly all mycobacterial species are free-living environmental saprophytes. A few, such as *Mycobacterium tuberculosis*, have evolved to cause transmissible human infection and eventually to become obligate human pathogens. The recent emergence and global spread of virulent clones of the environmental nontuberculous mycobacterium *M. abscessus* has provided a unique opportunity to examine the pathogenic evolution of mycobacteria.

RATIONALE: *M. abscessus*, a multidrug-resistant species of nontuberculous mycobacteria, has recently emerged as a major threat to individuals with cystic fibrosis (CF) and other chronic lung conditions. Infection rates within the CF community are increasing globally, driven in part by indirect person-to-person transmission of *M. abscessus*.

Currently more than 70% of infections in CF patients are caused by genetically clustered (and thus transmitted) isolates, of which the majority are from three dominant circulating clones (DCCs) that have emerged within the past 50 years and have spread globally. These clustered isolates are more virulent when tested in vitro and in vivo and result in worse clinical outcomes, which suggests that they are evolving from environmental saprophytes into obligate lung pathogens. We reasoned that functional genomic analysis of *M. abscessus* might identify important generalizable steps in this evolutionary trajectory and highlight potential interventions to mitigate this process for this and other emergent mycobacterial pathogens.

RESULTS: We initially sought to understand how the DCCs may have emerged. Using graphical pan-genome analysis, we found that horizontal gene transfer—particularly gene transfer of global transcriptional regulators—can provide an important mechanism for creating large phenotypic variance in environmental *M. abscessus* isolates, consequently enabling saltational evolution toward enhanced human infectivity. This process may be generalizable across mycobacterial species where gene gain or loss events have been associated with the pathogenic evolution of virulent clones in several species, including cluster 1a within *M. avium*, clone A within *M. canettii*, and the monophyletic *M. tuberculosis* complex from an *M. canettii*–like ancestor.
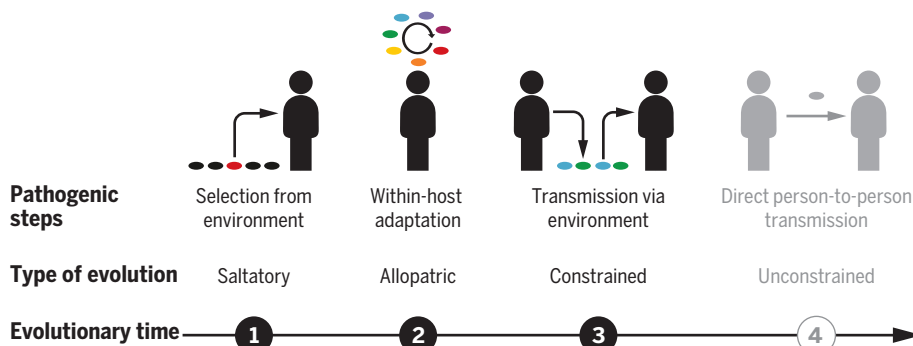
We next examined whether ongoing adaptation of infecting *M. abscessus* clones could further promote pathogenicity. We reconstructed the evolutionary trajectories of individual *M. abscessus* subclones within chronically infected patients, identifying convergent allopatric evolution within and between individuals as a key driver for pathogenic adaptation. Specifically, recurrent mutations within a small set of genes from a single functional network are likely to drive enhanced macrophage survival and increased virulence in vivo.

We observed reduced transmission fitness for many adaptive mutations. For two frequently mutated genes (*phoR* and the GPL locus), knockouts showed impaired survival on fomites. These findings suggest that within-host evolution is constrained, while an environmental intermediary is required for transmission.

CONCLUSION: Our results point to a generalizable model for mycobacterial pathogenic evolution. Initially, horizontal gene acquisition by environmental clones (particularly of genes with global transcriptional effects) drives saltational evolution and increases virulence, giving rise to the ancestors of the dominant circulating clones of *M. abscessus* and driving the emergence of virulent clones in other mycobacterial species.

Next, within-host adaptation during chronic infection drives increased intracellular survival within macrophages and inflammatory lung damage. However, pathogenic evolution is constrained while *M. abscessus* is transmitted through environmental intermediaries, because the most highly adapted strains lose transmission fitness through reduced fomite survival.

Ultimately, we predict that opportunities for direct transmission of emergent mycobacteria (potentially through increases in population density and/or host susceptibility) will permit unconstrained, accelerated evolution into an obligate human pathogen (as occurred in *M. tuberculosis* several thousand years ago). Our findings indicate how key interventions, such as early treatment and cross-infection control, might restrict existing pathogens and prevent new, emergent ones. ∎

Pathogenic steps	Selection from environment	Within-host adaptation	Transmission via environment	Direct person-to-person transmission
Type of evolution	Saltatory	Allopatric	Constrained	Unconstrained
Evolutionary time	1	2	3	4

Steps involved in mycobacterial pathogenic evolution. (1) Horizontal gene acquisition by environmental clones drives saltational evolution and gives rise to the ancestors of the dominant circulating clones of *M. abscessus* (and virulent clones within other mycobacterial species). (2) Allopatric within-host adaptation during chronic infection drives increased intracellular survival within macrophages and inflammatory lung damage. (3) Evolution is constrained while *M. abscessus* is transmitted through environmental intermediaries because the most highly adapted strains lose transmission fitness through reduced fomite survival. (4) Opportunities for direct transmission of emergent mycobacteria (potentially through increases in population density and/or host susceptibility) permit unconstrained, accelerated evolution (as occurred in *M. tuberculosis*).

S READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abb8699

PALEOECOLOGY

Widespread reforestation before European influence on Amazonia

M. B. Bush[1]*, M. N. Nascimento[1,2], C. M. Åkesson[1]†, G. M. Cárdenes-Sandí[3], S. Y. Maezumi[2], H. Behling[4], A. Correa-Metrio[5], W. Church[6], S. N. Huisman[2], T. Kelly[7], F. E. Mayle[8], C. N. H. McMichael[2]*

An estimated 90 to 95% of Indigenous people in Amazonia died after European contact. This population collapse is postulated to have caused decreases in atmospheric carbon dioxide concentrations at around 1610 CE, as a result of a wave of land abandonment in the wake of disease, slavery, and warfare, whereby the attendant reversion to forest substantially increased terrestrial carbon sequestration. On the basis of 39 Amazonian fossil pollen records, we show that there was no synchronous reforestation event associated with such an atmospheric carbon dioxide response after European arrival in Amazonia. Instead, we find that, at most sites, land abandonment and forest regrowth began about 300 to 600 years before European arrival. Pre-European pandemics, social strife, or environmental change may have contributed to these early site abandonments and ecological shifts.

The scale and the spatial and temporal patterns of human population dynamics in Amazonia have long been controversial. Early models suggest an exponential increase that perhaps continued until European contact (*1*), but more recent assessments suggest that population growth was slowing by around 1200 CE (hereafter, if not specified, years are of the Common Era) (*2*), perhaps nearing a carrying capacity. After 1492, an estimated 90 to 95% of the Indigenous population was lost to waves of disease—including smallpox, influenza, measles, and the common cold—sweeping through "virgin soil" communities or to warfare and slavery (*3*). This catastrophic loss of life resulting from European colonization has been called the "Great Dying" of the Indigenous peoples of the Americas (*4*). The population collapse is commonly considered a turning point in human influence on Amazonian landscapes, as inhabited sites were abandoned, and land used for crop cultivation became fallowed (Fig. 1). One suggested manifestation of the

attendant surge in forest regrowth was a 7 to 10 parts per million (ppm) drop in atmospheric CO_2 concentrations known as the Orbis spike, which began around 1610 and presaged relatively low concentrations until around 1750 (Fig. 1) (*4*)—a decline that deepened the cool-

ing of the Little Ice Age (1400 to 1800) (*5*). Whether the scale of CO_2 variation forming the Orbis spike was truly an unusual event has been questioned (*6*), as has the link to New World depopulation (*7*). Of all the Americas, the greatest potential carbon response to the Great Dying would have been manifested in the vast, high-biomass forests of Amazonia (*8*, *9*). If the decline in CO_2 concentrations was caused by the Great Dying, the depopulation and reforestation must have been rapid and widespread (Fig. 1).

Eyewitness accounts of the state of Amazonian populations in the first 200 years after European contact are sparse, but three accounts stand out: those of Carvajal (*10*), Acuña (*11*), and La Condamine (*12*). Lured by rumors of gold, the first large Spanish expedition entered lowland Amazonia in 1541 (*10*). Friar Carvajal, the diarist of Orellana's expedition, reported large, healthy populations along many portions of the river, with no suggestion of mass disease (*10*). A similarly positive account of societies and living conditions is provided by Acuña, a Spanish priest who traveled from Quito to Belem in 1639. These two early accounts could have been flavored by a desire to present a land of opportunity to royal courts in Europe (*12*). In contrast, a French surveyor, La Condamine, traveled down the Amazon River
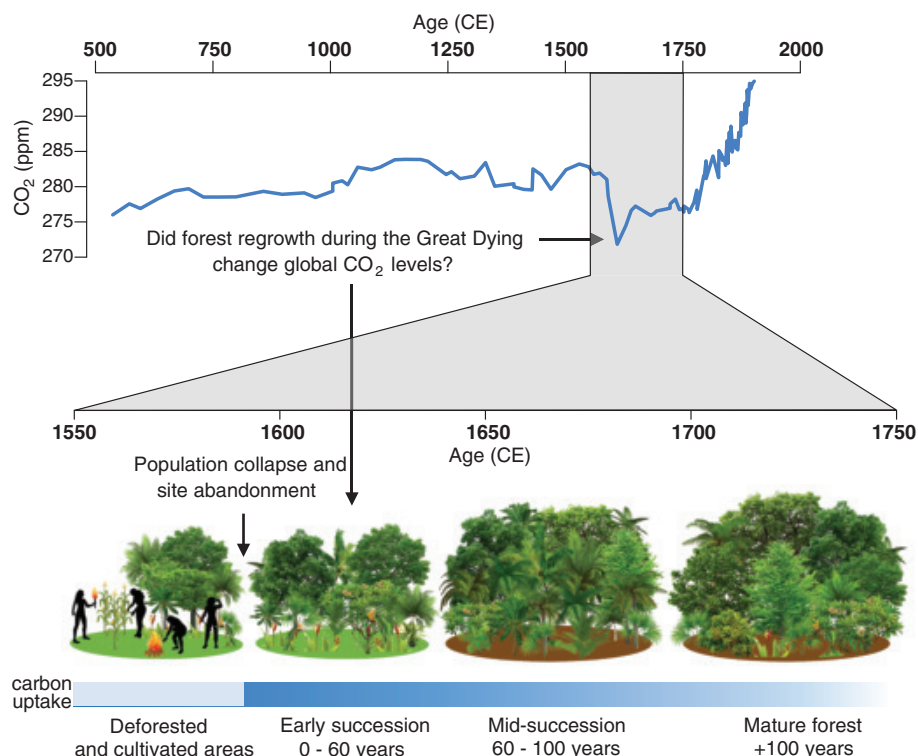
[1]Institute for Global Ecology, Florida Institute of Technology, Melbourne, FL, USA. [2]Department of Ecosystem and Landscape Dynamics, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, Netherlands. [3]Escuela Centroamericana de Geología, University of Costa Rica, San José, Costa Rica. [4]Department of Palynology and Climate Dynamics, University of Goettingen, Goettingen, Germany. [5]Instituto de Geología, Universidad Nacional Autónoma de México, Ciudad de México, México. [6]Department of Earth and Space Sciences, Columbus State University, Columbus, GA, USA. [7]School of Geography, Queen Mary University of London, Mile End, London, UK. [8]Department of Geography and Environmental Science, University of Reading, Whiteknights, Reading, Berkshire, UK.
†Present address: School of Geography and Sustainable Development, University of St. Andrews, St. Andrews, UK.
*Corresponding author. Email: mbush@fit.edu (M.B.B.); c.n.h.mcmichael@uva.nl (C.N.H.M.)

Fig. 1. How reforestation could relate to atmospheric carbon concentrations. Atmospheric CO_2 concentrations (blue line) (*5*) showing the 7 to 10 ppm decline at 1610 are attributed to the Great Dying. Expected carbon uptake patterns associated with occupation, abandonment, and forest recovery are shown for 1550 to 1750, assuming that the most intense sequestration takes place in the first century of succession (*41*).

in 1743 but did not record the same high density of people on the riverbanks, suggesting a partially depopulated landscape (*12*). Although it seems most probable that any vegetation change associated with the Great Dying took place in Amazonia after 1639 (the year of Acuña's trip), we investigate the possibility that it occurred between 1550 and 1750, concurrent with low CO_2 levels (*8*, *9*).

Although it is challenging to estimate pre-collapse population size, assumptions of near-synchronous forest regrowth are readily testable using paleoecology. Fossil pollen recovered from lake sediments provides a metric for reconstructing local forest cover and land use (*13*). If the Great Dying induced the rapid, synchronous forest regrowth, then pollen contained in these fossil pollen records should show the strongest switch of the past 2000 years between open ground and weedy species (signaling a deforested landscape) to dominance of forest taxa between 1550 and 1750 (Fig. 1). This signal should be particularly strong, because lakes were preferred settlement sites for Indigenous populations (*13*).

Although there is not a 1:1 relationship between pollen percentages and forest cover percentages, records from dense rainforest settings do provide a sensitive index of even small-scale disturbances (*14*, *15*). Consistently, within the forested portion of Amazonia, relatively undisturbed forest produces 95 to 100% forest pollen (*14*). Where human disturbance occurs, forest pollen percentages decrease, while percentages of open ground and shrubby taxa increase. Fossil pollen of weeds, grasses, and crops is usually accompanied by charcoal, which is a direct indicator of anthropogenic forest burning, as natural (lightning-induced) fires rarely occur in Amazonian forests (*16*). An important taxon in assessing forest disturbance and early stages of recovery is *Cecropia,* a short-lived, fast-growing, shade-intolerant pioneer tree that produces abundant and easily identifiable pollen (*17*). *Cecropia* occurs naturally as a canopy gap–colonist in forests and hence is part of our forest pollen component, but it is favored by anthropogenic disturbance and commonly forms a dominant stage in forest succession on abandoned land (*18*). Abandoned fields in much of Amazonia would be expected to pass through a *Cecropia*-rich early successional stage. As the forest matures, *Cecropia* would be competitively excluded from all but forest gaps, and so a peak of *Cecropia* pollen in fossil pollen records should be a sensitive marker of the onset of the Great Dying. To determine whether reforestation between 1550 and 1750 was associated with changes in global CO_2 levels, we assess changes in land use, forest cover, and *Cecropia* abundances for the past 2000 years using fossil pollen and charcoal data derived from 39 lake sites across Amazonia (*19*).
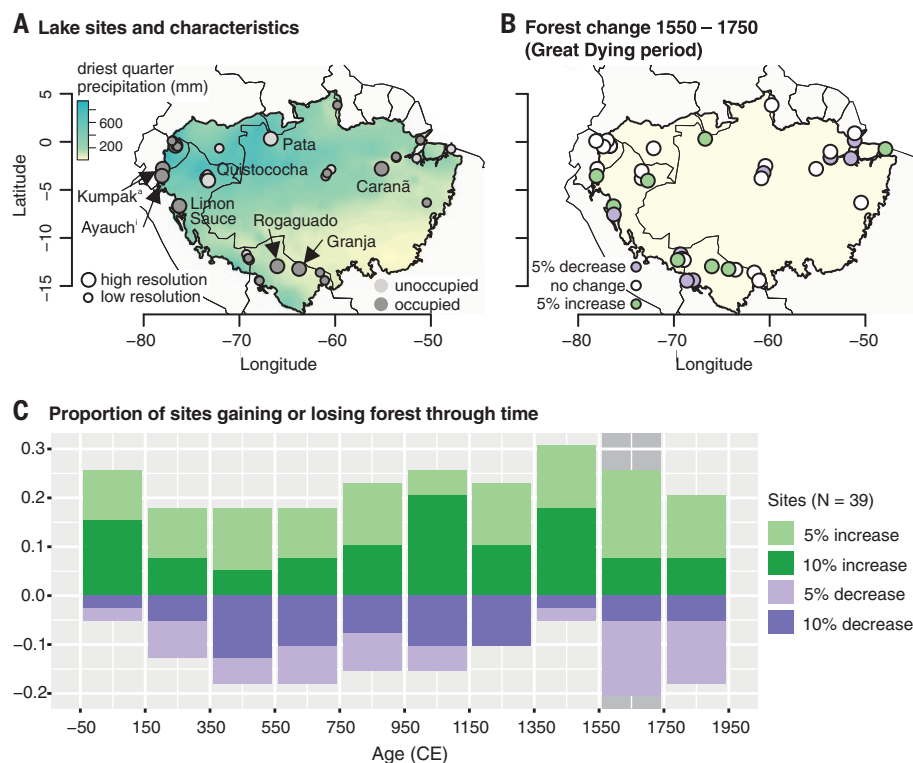

Fig. 2. Sites used in the analysis. (**A**) Distribution of 39 lake sediment records shown in the context of precipitation of the driest quarter (i.e., the three consecutive months per year with the lowest precipitation values) (*20*), with larger symbols representing the high-resolution sites (Figs. 3 and 4). Color coding indicates whether evidence of past human occupation was present in the record. (**B**) Forest change during the Great Dying period. Color coding indicates whether pollen percentages at each site had increased (green) or decreased (purple) at least 5% during the Great Dying period, compared with samples in the previous time window. (**C**) The proportion of sites that were losing or gaining either 5 or 10% of forest pollen (compared with samples in the previous time window) over the past 2000 years, using 200-year time bins (50 BCE to 1950 CE). Dark gray shading indicates the Great Dying period.

About 80% of the 39 sites contained signals of forest opening, burning, or cultivation consistent with pre-European occupation (Fig. 2A and data S1). A spatially and temporally heterogeneous pattern of deforestation, reforestation, and, by inference, carbon uptake was evident in the pollen data across the 39 sites (Fig. 2 and figs. S1 and S2). To assess even minor changes in forest cover, the proportion of sites showing a 5 or 10% increase or decrease in forest pollen was assessed. Sites exhibiting evidence of deforestation at values >5% peaked between 350 and 750 CE, whereas the proportion of sites showing evidence of reforestation was greatest between 750 and 1550 CE (Fig. 2C and figs. S1 and S2). During the Great Dying period, the number of sites where forest pollen was increasing roughly equaled those where it was falling in abundance (Fig. 2, B and C, and fig. S1), effectively rejecting the hypothesis of widespread and synchronous reforestation sufficient to cause decreases in atmospheric CO_2 levels. Instead of a strong signal of reforestation during or after the Great Dying, our empirical data show that changes in Amazonian land use and forest cover took place several centuries before European arrival (Fig. 2C).

Nine sites out of 39 contain at least 10 pollen samples from the past 1000 years (fig. S3) and provide centennial-scale temporal resolution (hereafter referred to as high-resolution sites). These sites offer the opportunity to investigate trajectories of site disturbance and forest recovery (Fig. 3 and data S1). Eight of the nine high-resolution sites contain pollen and charcoal evidence of occupation in the pre-Columbian era and provide evidence of changes in forest pollen percentages before, during, or after the Great Dying (Fig. 3 and data S1). Together, the nine sites do not show a pattern of synchronous or substantial reforestation during the Great Dying period (Fig. 3). Instead, these records contain a range of temporal patterns of pre-Columbian occupancy, i.e., long-term sustained use, intermittent use, and nonuse (*16–24*). An example of a site showing long-term use is Lake Caranã (Fig. 3), which provides evidence in its fossil pollen and charcoal record of a continuous history of occupation, with the frequent use of fire and maize (*Zea mays*) and squash (*Cucurbita* spp.)
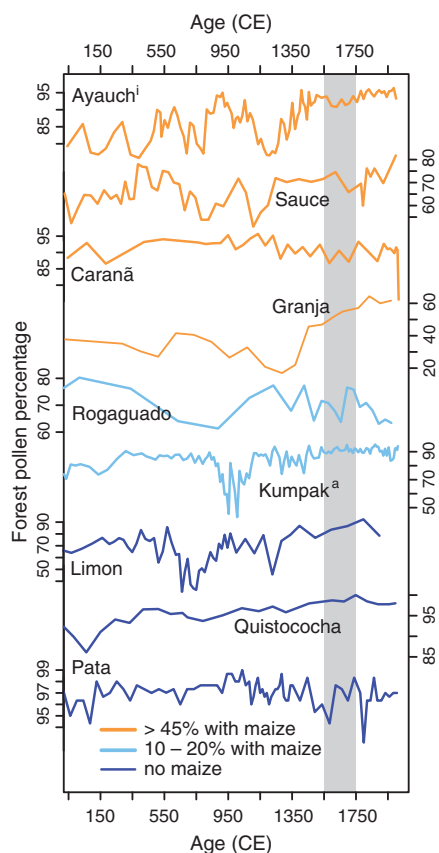
Fig. 3. Variability in pollen percentages of forest taxa over the past 2000 years, documented at the nine high-resolution lake sites. Sites are color coded by the percentage of samples in the record containing the presence of maize pollen, a direct indicator of cultivation (see tables S4 to S12 in data S1). Dark gray shading indicates the Great Dying period (1550 to 1750 CE).

Fig. 4. Changes in forest pollen abundance and forest burning. Percentage differences from the site mean (0 to 2000 CE) for (**A**) forest pollen and (**B**) *Cecropia*, for the nine high-resolution sites. Data are interpolated to 100-year time bins. (**C**) The charcoal index values represent the charcoal abundance data for each lake, scaled and standardized to values between 0 (absence of fire) and 100 (maximum charcoal abundance given that all samples contain charcoal). The standardization allows comparisons across sites (*15*). Mean values (black line) for (A) to (C) indicate the average value across all sites for each 100-year time window.

cultivation. Archaeological data at Caranã evidence the formation of Amazonian Dark Earth over the past 2000 years (*21*). Despite this intensity of use, no notable change in total forest pollen is associated with the Great Dying at Lake Caranã. Lakes Rogaguado and Granja were the only pollen records that showed an increase in forest pollen abundance by >10% between 1550 and 1750 (Fig. 3). Most of the increase at Rogaguado occurred between 1650 and 1750 CE. Although Granja lies in riparian forest today, it is within 1 km of flooded savanna. A modest increase in precipitation at this site could have caused an increase in forest pollen representation, but that same increase would not affect other locations that were already fully forested. Progressively wetter conditions at Granja over the past 2000 years should have led to increased forest cover, but instead, grassland was maintained by human activities until abandonment, at around 1300 to 1400 CE (*22*). After 1300 CE, forest cover increased from 20 to >40% in about a cen-
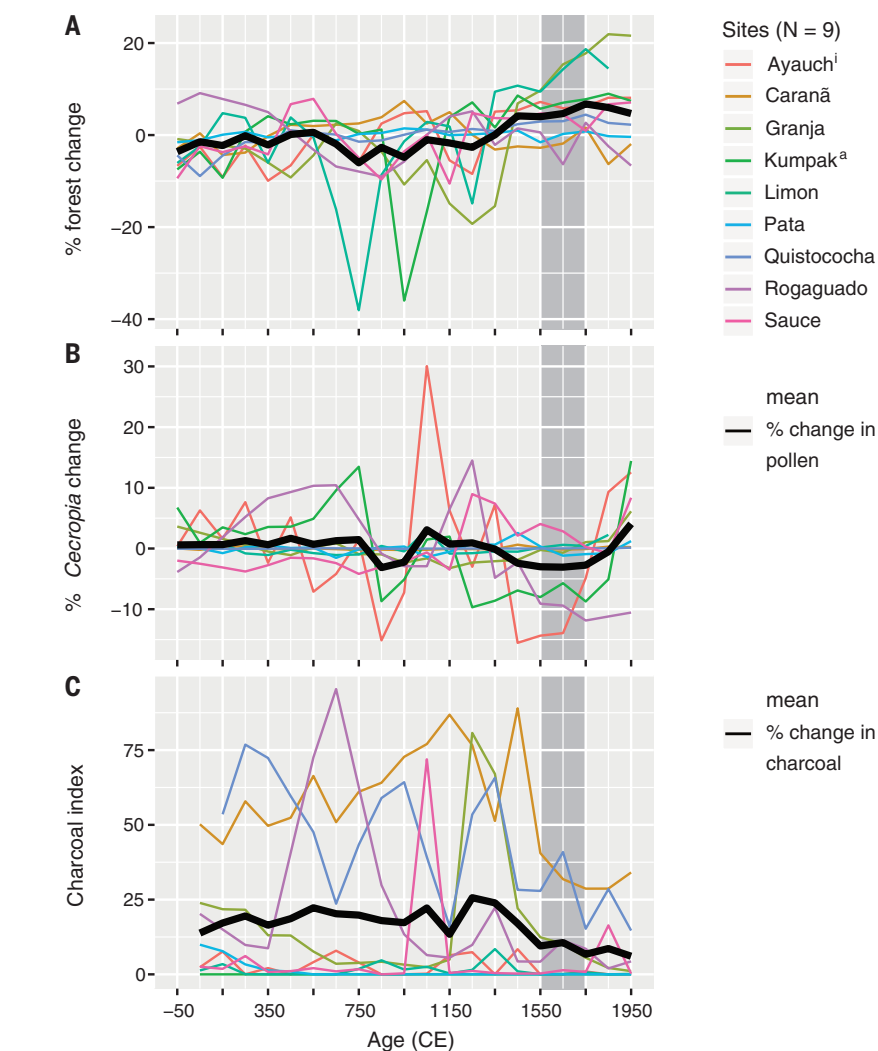
tury (Fig. 3). Thus, Granja was not a site where there was a cycle of deforestation, use, and reforestation, as seen at all the other occupied sites; rather, human activity suppressed a natural, climatically induced trend of rainforest expansion. Among the other sites containing pollen evidence of maize cultivation, fluctuations in forest pollen percentages, indicative of forest clearance and recovery, varied in timing. Sauce (*23*) and Ayauch[i] (*24*) had records of long-term forest clearance, whereas Lake Kumpak[a] (*25*) was used episodically. Quistococha (*26*) and the saline Lake Limon (*27*) provided evidence of human occupation (burning) but no crop pollen. None of the nine sites showed increases in forest pollen during the Great Dying. Lake Pata lies on an inselberg above the Amazon

plain, with very thin soils that preclude cultivation (*28*). A high-resolution pollen analysis of Lake Pata revealed no evidence of human occupancy and forest pollen percentages that were >95% throughout the past 2000 years (Fig. 3) (*28*).

When looking at deviations in forest pollen abundance from the 2000-year mean pollen percentage of each record, only Lake Rogaguado showed notable forest increases during the Great Dying period, particularly from 1650 to 1750 CE (Fig. 4A). Other lakes showed either no deviation throughout the record or forest percentages that increased relative to mean values between 950 and 1350, i.e., 300 to 600 years before the Great Dying event. The largest increase in the mean value of all forest pollen deviations occurred from

about 1250 to 1350, with no substantial increase during the Great Dying period (Fig. 4A, black line). Between 950 and 1350, *Cecropia* pollen percentages showed the largest increases (~5%) from mean values, whereas during the Great Dying, values were ~5% lower than the long-term mean (Fig. 4B). These data are inconsistent with a postulated peak in post-disturbance succession and rapid carbon uptake after the Great Dying. The negative deviations from the mean *Cecropia* values for each of the nine lakes correspond with charcoal declines. Such cessation of burning is strongly associated with land abandonment (Fig. 4, B and C). The discontinuous charcoal peaks at most sites suggest that forest burning was intermittent, and where it occurred, it ceased before the Great Dying (Fig. 4C)—a probable indicator of abandonment. The only site to show a decrease in burning between 1650 and 1750 was Lake Quistococha, which is located near the city of Iquitos, Peru, in a region where the Jesuits established the Mainas missions from 1638 to 1767 (*29*).

Our empirical observations are consistent with archaeologically derived models (*2*) that suggest stable or falling populations for centuries before European arrival. The mechanisms driving the cultural change or site abandonment from around 950 to 1350, however, have yet to be identified. We consider three possibilities that are not mutually exclusive: climate change, societal change, and disease. If climate were mainly responsible, lakes within close proximity to each other would be expected to contain synchronous changes in pollen and charcoal, and regional geographic trends associated with environmental gradients (e.g., Fig. 2A) should be evident in the dataset. But they are not. Trends in forest pollen abundance do not seem to have a distinct geographic pattern (Fig. 2B and figs. S1 and S2), and neighboring sites often show nonsynchronous peaks of fire and periods of fire absence (Fig. 4) (*30*, *31*), suggesting that climate change is unlikely to be the sole explanation for the abandonment of sites between 950 and 1250. Our findings do not, however, discount the possibility that climate change could have contributed to a societal response of changing land use and thus changing forest cover percentages (*23*). Isotopic data from regional speleothems and lake records both support a trend toward increasing climate variability between 800 and 1200 (*32*, *33*). De Souza *et al.* (*34*) suggested that complex, hierarchical societies relying on specialized food provisions were more susceptible to these climate variations than simpler, more egalitarian societies accessing a breadth of food resources. Many of the sites used in this study have no accompanying archaeological data to determine past human vulnerability to climate change. Nevertheless, even if human populations were resil-

ient, some sites may simply have become too flood- or drought-prone to continue to be desirable, forcing migration to new locations.

The timing of observed reforestation in the lowlands coincides with the relocation of an estimated 25% of the Indigenous population from the Andes into the coastal lowlands between about 1000 and 1200 CE (*35*). This migration is associated with the collapse of the Tiwanaku and Wari cultures, rapid climate change, and—as evidenced by cranial trauma—increased warfare (*36*, *37*). In the Amazon lowlands, increased hostility is inferred from archaeological contexts amid a cultural expansion reflected in the spread of the polychrome tradition of decorated pottery (*38*). To find abandonment of apparently unrelated lowland settings at the same time as that of the highlands raises the possibility of a common cause. Climate change, conflict, and disease could underlie both patterns of behavior. Disease outbreaks have yet to be documented in the lowlands, but skeletal remains provide evidence of the Andean expansion of tuberculosis between 1000 and 1300 (*39*). Trading between lowland and highland communities, which this period is known for (*40*), could have easily spread disease and created pre-European pandemics across the region. Thus, the interaction of climate change, social tensions, and possibly even the emergence of novel non-European diseases could have caused the observed destabilization of Amazonian populations centuries before the Great Dying period. Populations in some areas of Amazonia may already have been declining when Europeans arrived, a decline that was accelerated by the impacts of disease after European contact. Furthermore, our data suggest that the timing of reforestation was heterogeneous, with many sites showing an increase in forest cover as many as 600 years before the Great Dying. We find no evidence that human-induced vegetation change in Amazonia influenced global CO_2 concentrations either during the early reforestation event documented here or during the Great Dying.

REFERENCES AND NOTES

1. M. Arroyo-Kalin, *Archaeol. Int.* **20**, 122–136 (2018).
2. M. Arroyo-Kalin, P. Riris, *Philos. Trans. R. Soc. London Ser. B* **376**, 20190715 (2020).
3. N. D. Cook, *Born to Die: Disease and New World Conquest, 1492–1650* (New Approaches to the Americas Series, Cambridge Univ. Press, 1998).
4. S. L. Lewis, M. A. Maslin, *Nature* **519**, 171–180 (2015).
5. C. MacFarling Meure *et al.*, *Geophys. Res. Lett.* **33**, L14810 (2006).
6. E. Monnin *et al.*, *Science* **291**, 112–114 (2001).
7. J. Zalasiewicz *et al.*, *Anthropocene Rev.* **2**, 117–127 (2015).
8. A. Koch, C. Brierley, M. M. Maslin, S. L. Lewis, *Quat. Sci. Rev.* **207**, 13–36 (2019).
9. R. J. Nevle, D. K. Bird, W. F. Ruddiman, R. A. Dull, *Holocene* **21**, 853–864 (2011).
10. J. Medina, *The Discovery of the Amazon According to the Account of Friar Caspar de Carvajal and Other Documents*, H. C. Heaton, Ed. (American Geographical Society, 1934).
11. C. Acuña, *A Relation of the Great River of Amazons in South-America. Voyages and discoveries in South-America.* (S. Buckley, London, 1698); https://www.loc.gov/item/02009583/.
12. N. J. H. Smith, *The Amazon River Forest: A Natural History of Plants, Animals, and People* (Oxford Univ. Press, 1999).
13. M. B. Bush, M. R. Silman, *Front. Ecol. Environ.* **5**, 457–465 (2007).
14. B. S. Whitney *et al.*, *Holocene* **29**, 262–270 (2019).
15. M. B. Bush, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **177**, 5–17 (2002).
16. M. A. Cochrane, in *Tropical Fire Ecology*, M. A. Cochrane, Ed. (Springer Berlin Heidelberg, 2009), pp. 1–23.
17. J. C. I. Rodgers III, S. P. Horn, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **124**, 53–71 (1996).
18. R. C. Mesquita, K. Ickes, G. Ganade, G. B. Williamson, *J. Ecol.* **89**, 528–537 (2001).
19. Materials and methods are available as supplementary materials.
20. R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, A. Jarvis, *Int. J. Climatol.* **25**, 1965–1978 (2005).
21. S. Y. Maezumi *et al.*, *Nat. Plants* **4**, 540–547 (2018).
22. J. F. Carson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10497–10502 (2014).
23. M. B. Bush, A. Correa-Metrio, R. van Woesik, C. R. Shadik, C. N. H. McMichael, *Glob. Change Biol.* **23**, 3181–3192 (2017).
24. C. Åkesson, "Humans, vegetation and drought in Late-Holocene Amazonia," thesis, Florida Institute of Technology, Melbourne, FL (2019).
25. C. Åkesson *et al.*, *J. Ecol.* **109**, 432–446 (2021).
26. T. J. Kelly *et al.*, *J. Quaternary Sci.* **33**, 369–379 (2018).
27. G. M. Cárdenes-Sandí, "Reconstrucción del escenario ambiental durante los últimos la cuenca del Lago Limón, 2200 años de la zona de Amazonía Peruana," thesis, Universidad de Costa Rica (2011).
28. M. N. Nascimento *et al.*, *J. Biogeogr.* **46**, 2389–2406 (2019).
29. M.-E. Reeve, *Ethnohistory* **41**, 106–138 (1993).
30. M. B. Bush *et al.*, *Philos. Trans. R. Soc. London Ser. B* **362**, 209–218 (2007).
31. C. H. McMichael *et al.*, *Science* **336**, 1429–1431 (2012).
32. B. W. Bird *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8583–8588 (2011).
33. X. Wang *et al.*, *Nature* **541**, 204–207 (2017).
34. J. G. de Souza *et al.*, *Nat. Ecol. Evol.* **3**, 1007–1017 (2019).
35. L. Fehren-Schmitz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9443–9448 (2014).
36. E. Arkush, T. A. Tung, *J. Archaeol. Res.* **21**, 307–369 (2013).
37. C. R. Ortloff, A. L. Kolata, *J. Archaeol.* **20**, 195–221 (1993).
38. C. D. P. Moraes, E. G. Neves, *Amazônica* **4**, 122–148 (2012).
39. E. A. Nelson, J. E. Buikstra, A. Herbig, T. A. Tung, K. I. Bos, *Int. J. Paleopathol.* **29**, 128–140 (2020).
40. W. B. Church, A. Von Hagen, in *The Handbook of South American Archaeology*, H. Silverman, W. H. Isbell, Eds. (Springer, 2008), pp. 903–926.
41. P. A. Martin, A. C. Newton, J. M. Bullock, *Proc. Biol. Sci.* **280**, 20132236 (2013).

ISLAND ECOLOGY

The human dimension of biodiversity changes on islands

Sandra Nogué[1]†*, Ana M. C. Santos[2,3,4,5], H. John B. Birks[6,7], Svante Björck[8], Alvaro Castilla-Beltrán[1], Simon Connor[9,10], Erik J. de Boer[11], Lea de Nascimento[12,13], Vivian A. Felde[6], José María Fernández-Palacios[12], Cynthia A. Froyd[14], Simon G. Haberle[9,10], Henry Hooghiemstra[15], Karl Ljung[8], Sietze J. Norder[16], Josep Peñuelas[17,18], Matthew Prebble[9,19], Janelle Stevenson[9,10], Robert J. Whittaker[20,21], Kathy J. Willis[22], Janet M. Wilmshurst[13,23], Manuel J. Steinbauer[24,25]†*

Islands are among the last regions on Earth settled and transformed by human activities, and they provide replicated model systems for analysis of how people affect ecological functions. By analyzing 27 representative fossil pollen sequences encompassing the past 5000 years from islands globally, we quantified the rates of vegetation compositional change before and after human arrival. After human arrival, rates of turnover accelerate by a median factor of 11, with faster rates on islands colonized in the past 1500 years than for those colonized earlier. This global anthropogenic acceleration in turnover suggests that islands are on trajectories of continuing change. Strategies for biodiversity conservation and ecosystem restoration must acknowledge the long duration of human impacts and the degree to which ecological changes today differ from prehuman dynamics.

G lobally, human activities dominate ecological systems (1, 2) and are considered the main drivers for accelerating contemporary ecosystem transformation (3–6). The pressing need to evaluate the extent and dimensions of human impacts and the desire to restore "wild" systems have sparked controversy concerning the value of establishing prehuman baselines (7–9) and about the nature and timing of the onset of the Anthropocene (10–12). Archaeological and other paleodata on human impacts in continental systems reveal an increasingly human-transformed planet intensifying around the end of the Pleistocene (2, 13, 14). The lengthy time frame of human modification of ecosystem dynamics in continental contexts, spanning periods of substantial postglacial climate change, complicates the definition of prehuman baselines and hinders the investigation of natural ecosystem processes (15, 16).

In contrast to continents, most remote oceanic islands were colonized by people relatively recently, within the past 3000 years, when climates were similar to present conditions (17). The recent nature of human settlement means that the archaeological, paleoecological, and climate records are often more precisely resolved on well-studied islands compared with continents and are potentially more relevant for understanding remnant ecosystems and informing conservation and ecosystem restoration agendas. Therefore, island ecosystems provide opportunities to quantify the critical ecological transition from prehuman to human-dominated ecosystems (4, 15) and allow anthropogenic impacts on ecosystems to be placed within the context of long-term prehuman ecological dynamics (16–20). Although numerous studies have documented the timing, waves, and processes of species extinctions that accompanied human arrival on islands (18–24), paleoecological data networks now allow systematic quantification of ecosystem transformations on islands globally. Here, we analyzed fossil pollen time series for multiple independent islands from all the major archipelagos and oceans and across latitudes using

a breakpoint regression approach to test for altered rates and directionality of pollen, and thus vegetation compositional turnover, connected with human colonization (25) within an overall time frame of the past 5000 years. These time series of millennial-scale dynamics allow the assessment of whether the rates of vegetation compositional change consistently accelerated across multiple islands after initial human arrival. Our method uses ordination analyses to characterize the major gradient of compositional variation in the pollen data for each island, quantifying the mean rate of change through time before and after human arrival (Fig. 1), thereby allowing us to assess how human populations affected islands differently from natural perturbations (23).

Our results show that human arrival systematically accelerated directional compositional change in island ecosystems (Figs. 1 and 2). Rates of pollen compositional turnover increased after human arrival by up to a factor of 11, with large differences among islands (i.e., a median of 10.7 times higher turnover after human arrival, with a mean of 20.8 ± 26.5 times higher turnover). This acceleration is a globally consistent pattern observed on 24 of 27 islands independently of current and past island area, latitude, isolation, and elevation of the sampling site [Fig. 3, B to G, and tables S3 and S4 (25)]. Islands that were settled more recently, such as the Poor Knights archipelago in New Zealand (13th century) (19) and the Galápagos Islands (16th century) (26), show a steeper increase in the rate of turnover change ($P = 0.008$, $R^2 = 0.22$; linear regression with log-transformed arrival time; Fig. 3A) than on islands where humans arrived >1500 years ago [e.g., New Caledonia (27) and Fiji (28)]. This indicates either that the islands settled earlier were more resilient to human arrival or, more likely, that the recent major compositional turnover observed is explained by introduced species, land use practices, and technology deployed by later settlers being more transformative than those of earlier settlers. In addition, those islands colonized

>3000 years ago appear to show some declines in rates of compositional turnover toward the end of the sequence, although there are too few cases (n = 5) to draw firm conclusions.

For many islands, the model implementing a prescribed breakpoint at the time of human arrival closely fits the observed patterns in compositional turnover (Fig. 1). Human arrival estimates fall within the 95% confidence intervals of the optimal breakpoints (representing the greatest change in turnover in each record)

for 41% of islands. Human arrival times are within 500 years of the optimal breakpoint for 70% of islands and within 1000 years for 81% of islands (median 329 years compared with 953 for randomized data simulations; table S5 and Fig. 2). There is no tendency for optimized breakpoints to be systematically earlier or later than estimated human arrival time (t test with null model of mean difference being 0, P = 0.27). A systematic difference would have either indicated earlier human arrival or de-

layed human impact. On some islands, initial human arrival is not associated with a major shift in turnover [Fig. 1 and fig. S1 (25)]. These results might reflect the specific local characteristics of the study site. For example, on La Gomera (Canary Islands), the sedimentary sequence was collected at an elevation of 1250 m above sea level in one of the largest remnant areas of laurel forest, where paleoecological analyses showed no evidence of human impacts (29). On other islands, e.g., Hispaniola,
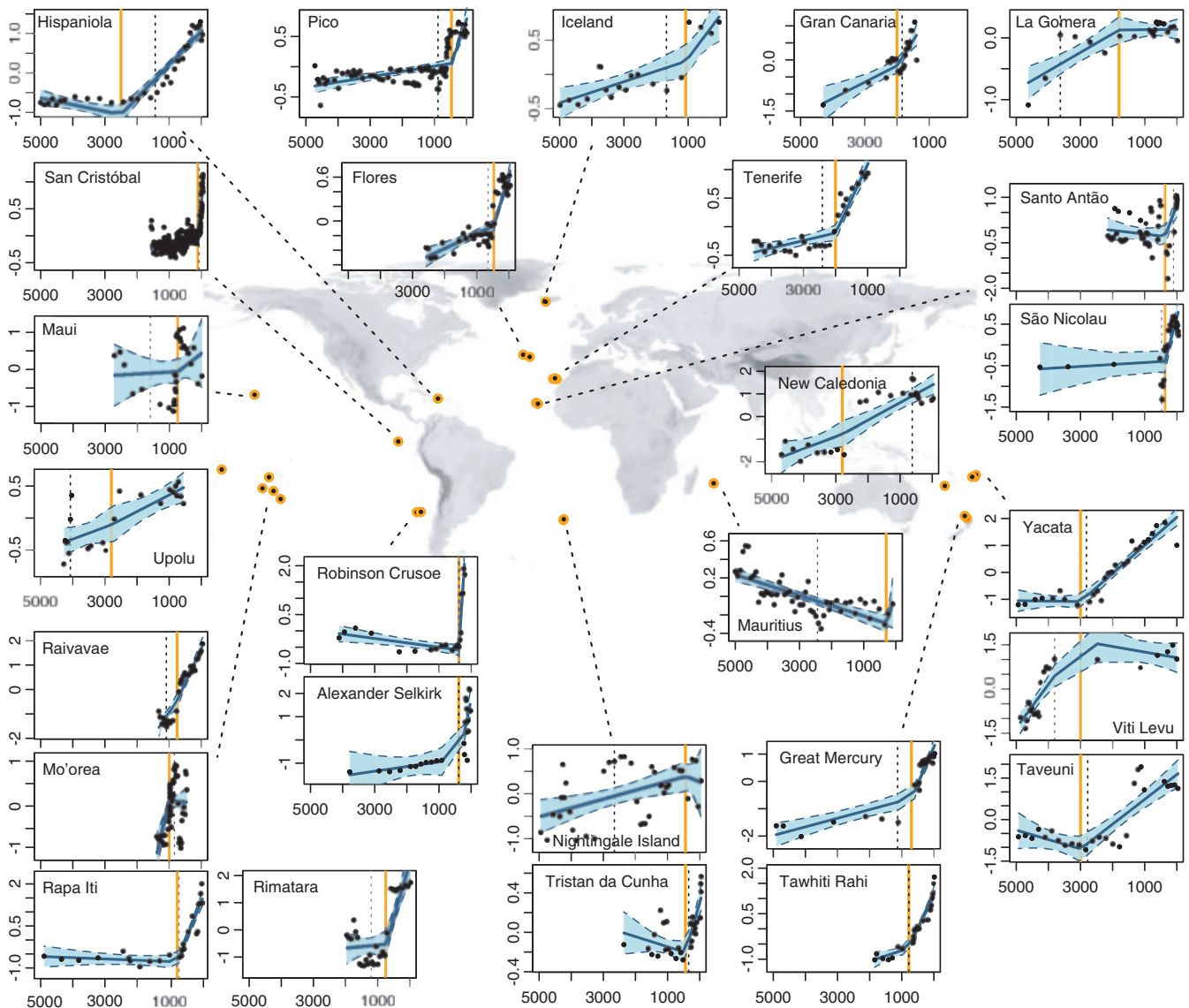
Fig. 1. Human arrival accelerated compositional turnover on islands. Global analysis of rate of palynological and thus vegetation compositional turnover (slope of the line) for 27 representative fossil pollen records from sedimentary sequences on islands. The x-axes represent calibrated years before present (i.e., years before 1950) calculated using Bayesian age-depth models for each island (25). The y-axes represent the major gradient in pollen composition quantified by the ordination axis 1 scores of separate detrended correspondence analyses (DCAs) of each sequence. The units are measured in DCA axis scores, which approximate the SD_{ptt}, with an SD of 4,

corresponding roughly to 100% compositional turnover. These plots show results of breakpoint analyses of the rate of compositional turnover with the date of recorded human arrival as the prescribed breakpoint. The recorded date of human arrival is indicated by the vertical orange lines (see table S3 for details). Scaling varies among panels. Shaded areas (blue) depict 95% confidence intervals of the models. A second continuous breakpoint analysis detecting the major statistical change point in turnover rate intrinsic was applied to the data. This "optimized breakpoint" is indicated by the vertical dashed black lines.

shifts in vegetation turnover differed from the time of human arrival, as estimated based on archaeological or historical sources, suggesting a lag between colonization and onset of significant directional change (Fig. 1).

Our analysis also shows that ecological change is an integral part of island systems, with changes observed both before recorded human arrival [directional change in composition measured in standard deviations of pollen taxon turnover (SD_{ptt}) per 100 years:

median turnover 1.7×10^{-2} SD_{ptt}/100 years and mean turnover $4.0 \pm 6.8 \times 10^{-2}$ SD_{ptt}/100 years] and after human arrival (median turnover 14.7×10^{-2} SD_{ptt}/100 years and mean turnover $23.3 \pm 29.8 \times 10^{-2}$ SD_{ptt}/100 years)] (Fig. 2). Results show that the rate of directional turnover before human arrival was slower, in contrast to human agencies of change. Natural drivers of ecosystem change on islands, operating before and alongside humans, include volcanic activities, fire, climate change (episodes such

as the "Little Ice Age"), earthquakes, extreme weather events (e.g., droughts and cyclones), and sea-level fluctuations (20, 30, 31). Although not measurable with the precision necessary to include formally within our analysis, volcanic activities and natural climate fluctuations have likely not increased over the analyzed time frame across the islands studied and thus cannot explain the systematic increase and varied timing of directional turnover observable across islands (25). Climate warming in the past 50 years, by contrast, is too recent to be detectable within our dataset. Over the time frame of the past 5000 years, direct human impacts greatly outweighed other processes that shaped island biodiversity and species interactions (32, 33).

Moreover, ecological legacies of human arrival on islands may persist for centuries and are often irreversible. An example is Tawhiti Rahi in the Poor Knights archipelago, which is currently uninhabited (19). Immediately after initial arrival by Polynesians in the 13th century, the island's forest cover was cleared by fire for human habitation and gardens. After a massacre of local Ngatiwai inhabitants on Tawhiti Rahi in 1820, local *kaitiaki* (guardians) declared the islands *wahi tapu* (protected by a sacred covenant), after which time there was no subsequent settlement. Despite the island becoming totally reforested within 150 years, the current forest composition is completely

Fig. 2. Rates of turnover before and after human arrival. Change in the rate of pollen compositional turnover before (left) and after recorded date of human settlement (purple) for the time series of fossil pollen records for each of 27 islands globally, where each island's sequence has been subject to a separate ordination analysis using DCA. Rate of pollen taxon turnover is quantified as the absolute slope in the relationship between ordination scores of the first axis of each DCA with time. The units approximate SD_{ptt}/100 years. The presettlement rate of compositional turnover is represented on the left (median: 1.7×10^{-4}; mean: 4.0×10^{-4}) and the rate after human arrival is represented on the right (median: 14.7×10^{-4}; mean: 23.3×10^{-4}). The difference is highly significant ($P < 0.004$; paired t-test). See (21) for details.
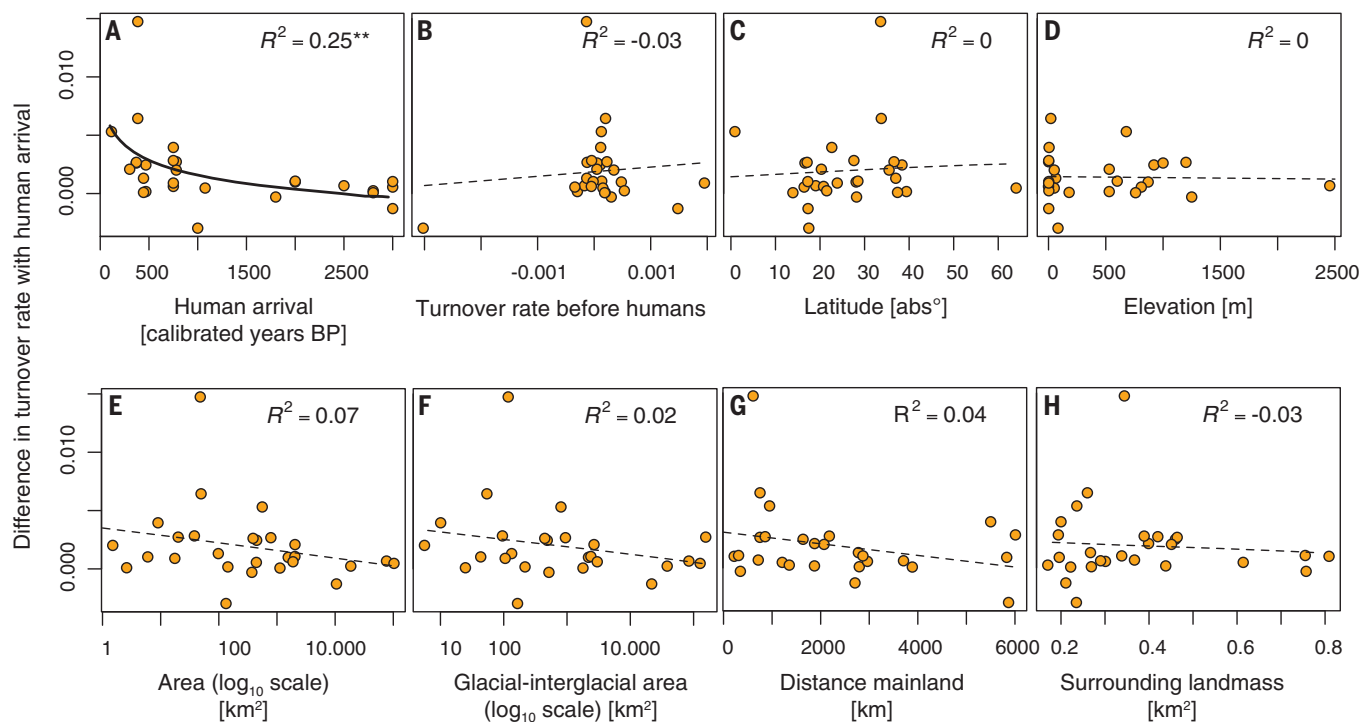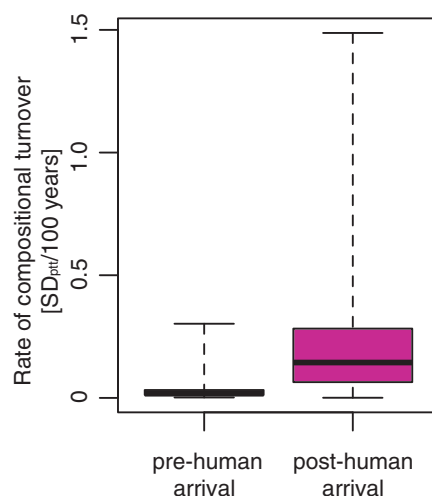
Fig. 3. Differences between the prehuman and human-dominated turnover scale with human arrival times. Relationships between the change in the rate of pollen compositional turnover before and after human arrival and several island features showing a curvilinear decrease in observed turnover as the time elapsed since the first colonization increases (**A**), but no relationship with turnover rate before human arrival (**B**), latitude (**C**), elevation of the coring site (**D**), island area (**E**), glacial-interglacial area (**F**), isolation (represented by distance to mainland) (**G**), or surrounding landmass (**H**). **$P < 0.01$ (A).

different from that of the prehuman period. In contrast to the Poor Knights archipelago, most currently inhabited islands have experienced at least two distinct waves of settlement, each having distinctive signatures of change and leaving increasingly complex legacies (24, 30).

Preparing and managing for ecosystem change is one of the major challenges that island societies currently face as islands experience continued or accelerated threats from detrimental land use practices (12), new species invasions (24, 34), sea-level rise (35), and climate change (11, 17), in addition to naturally occurring disturbances. The challenges are made more difficult because these processes are affecting native ecosystems where vegetation communities have already been severely degraded or lost, species have gone extinct (15, 21), and important mutualistic plant–animal interactions have been disrupted (36). Our results show little indication that these human-affected ecosystems are either similar to or returning to the dynamic baselines observed before human arrival. Therefore, anthropogenic impacts on islands are lasting components of these systems typically involving initial clearance (e.g., using fire) and are compounded by the introduction of a range of introduced species and extinctions of endemic species and ongoing disturbances. This contrasts with turnover after natural disturbances in the prehuman period, when island ecosystems often recovered rapidly to predisturbance states [e.g, (20, 31)]. Whereas for many islands, widescale return to precolonization ecosystems is an unrealistic goal, paleoecological data such as those analyzed here may serve to inform targeted ecosystem restoration efforts within islands, providing insights into previous system states and their responsiveness to global change processes (9, 37).

REFERENCES AND NOTES

1. P. M. Vitousek, H. A. Mooney, J. Lubchenco, J. M. Melillo, *Science* **277**, 494–499 (1997).
2. L. Stephens *et al.*, *Science* **365**, 897–902 (2019).
3. S. E. Connor *et al.*, *J. Biogeogr.* **39**, 1007–1023 (2012).
4. S. Nogué *et al.*, *Nat. Ecol. Evol.* **1**, 181 (2017).
5. W. Steffen, W. Broadgate, L. Deutsch, O. Gaffney, C. Ludwig, *Anthropocene Rev.* **2**, 81–98 (2015).
6. M. J. Steinbauer *et al.*, *Nature* **556**, 231–234 (2018).
7. K. J. Willis, R. M. Bailey, S. A. Bhagwat, H. J. B. Birks, *Trends Ecol. Evol.* **25**, 583–591 (2010).
8. R. J. Hobbs *et al.*, *Glob. Ecol. Biogeogr.* **15**, 1–7 (2006).
9. A. D. Barnosky *et al.*, *Science* **355**, eaah4787 (2017).
10. W. F. Ruddiman, *Prog. Phys. Geogr.* **42**, 451–461 (2018).
11. J. Zalasiewicz *et al.*, *Prog. Phys. Geogr.* **43**, 319–333 (2019).
12. C. S. M. Turney *et al.*, *Sci. Rep.* **8**, 3293 (2018).
13. E. C. Ellis, D. Q. Fuller, J. O. Kaplan, W. G. Lutters, *Elementa (Washington, DC)* **1**, 000018 (2013).
14. Y. Malhi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 838–846 (2016).
15. H. J. B. Birks, *Plant Ecol. Divers.* **12**, 189–385 (2019).
16. J. R. Wood, G. L. W. Perry, J. M. Wilmshurst, *Funct. Ecol.* **31**, 1012–1020 (2017).
17. C. Nolan *et al.*, *Science* **361**, 920–923 (2018).
18. W. D. Gosling *et al.*, *J. Biogeogr.* **47**, 600–614 (2020).
19. J. M. Wilmshurst *et al.*, *Conserv. Biol.* **28**, 202–212 (2014).
20. J. M. Wilmshurst, M. S. McGlone, T. R. Partridge, *N. Z. J. Bot.* **35**, 79–96 (1997).
21. D. W. Steadman, *Science* **267**, 1123–1131 (1995).
22. D. A. Burney, T. F. Flannery, *Trends Ecol. Evol.* **20**, 395–401 (2005).
23. R. J. Whittaker, J. M. Fernández-Palacios, *Island Biogeography: Ecology, Evolution, and Conservation* (Oxford Univ. Press, 2007).
24. J. R. Wood *et al.*, *Environ. Conserv.* **44**, 348–358 (2017).
25. Materials and methods are available as supplementary materials.
26. A. Restrepo *et al.*, *Ecology* **93**, 1853–1866 (2012).
27. J. Stevenson, R. Dodson, I. P. Prosser, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **168**, 97–123 (2001).
28. G. Hope, J. Stevenson, W. Southern, in *The Early Prehistory of Fiji*, G. Clark, Ed. (ANU Press, Terra Australis Series 31, 2009), pp. 63–86.
29. S. Nogué, L. de Nascimento, J. M. Fernández-Palacios, R. J. Whittaker, K. J. Willis, *J. Ecol.* **101**, 368–377 (2013).
30. B. Rolett, J. Diamond, *Nature* **431**, 443–446 (2004).
31. J. M. Wilmshurst, M. S. McGlone, *Holocene* **6**, 399–411 (1996).
32. M. R. Helmus, D. L. Mahler, J. B. Losos, *Nature* **513**, 543–546 (2014).
33. H. Kreft, W. Jetz, J. Mutke, G. Kier, W. Barthlott, *Ecol. Lett.* **11**, 116–127 (2008).
34. D. Moser *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9270–9275 (2018).
35. S. J. Norder *et al.*, *Glob. Ecol. Biogeogr.* **27**, 500–505 (2018).
36. C. N. Kaiser-Bunbury, A. Traveset, D. M. Hansen, *Perspect. Plant Ecol. Evol. Syst.* **12**, 131–143 (2010).
37. J.-C. Svenning, *Am. J. Bot.* **105**, 963–966 (2018).

ECOSYSTEM ENGINEERS

Equids engineer desert water availability

Erick J. Lundgren[1,2,3]*, Daniel Ramp[1], Juliet C. Stromberg[4], Jianguo Wu[4,5], Nathan C. Nieto[6]†, Martin Sluk[7], Karla T. Moeller[4], Arian D. Wallach[1]

Megafauna play important roles in the biosphere, yet little is known about how they shape dryland ecosystems. We report on an overlooked form of ecosystem engineering by donkeys and horses. In the deserts of North America, digging of ≤2-meter wells to groundwater by feral equids increased the density of water features, reduced distances between waters, and, at times, provided the only water present. Vertebrate richness and activity were higher at equid wells than at adjacent dry sites, and, by mimicking flood disturbance, equid wells became nurseries for riparian trees. Our results suggest that equids, even those that are introduced or feral, are able to buffer water availability, which may increase resilience to ongoing human-caused aridification.

Terrestrial large herbivores (henceforth megafauna) have undergone extensive extinctions and range contractions beginning during the late Pleistocene (100,000 to 12,000 years before the present) and continuing today (1, 2). Although climate change at the end of the last glacial maxima may have played a contributing role (3), emerging consensus indicates that most prehistoric losses were driven by human activity (1, 2). In tropical and temperate ecosystems, megafauna declines are linked to the formation of closed woodlands, increased wildfire, and reduced dispersal of large-seeded plants (4). Less is known, however, about how megafauna may have shaped dryland ecosystems, which comprise a third of Earth's surface (5).

Water is the main limiting resource in dryland ecosystems. It determines species composition, food web structure, and vegetation dynamics (6, 7). Yet, the capacity for animals to enhance water availability by exposing subsurface water has received little attention. Wild donkeys (*Equus africanus asinus*) and horses (*E. ferus caballus*), as well as most other equids and all elephant species, regularly dig wells of up to 2 m in depth (Fig. 1, A to D; see table S1 for review). We evaluated well digging and its associated ecosystem effects in a North American system where equids have established feral populations.

We surveyed four Sonoran Desert groundwater-fed streams every 2 to 4 weeks over three summers (table S2). At each site we mapped "background" (e.g., already present) and "equid

well" water [data S1 and (8)]. Streams were 7 to 32 km apart and were ~300 to 1800 m long (table S2). Like many desert streams, site hydrology was highly variable, as was the relative contribution of equid wells. Equid wells were particularly important to provisioning water in midsummer as temperatures increased and water tables receded (Fig. 1E). At one fully intermittent stream that lost all background water, equid wells provided 100% of surface water. Even at sites which remained perennial (background water retained at headwater
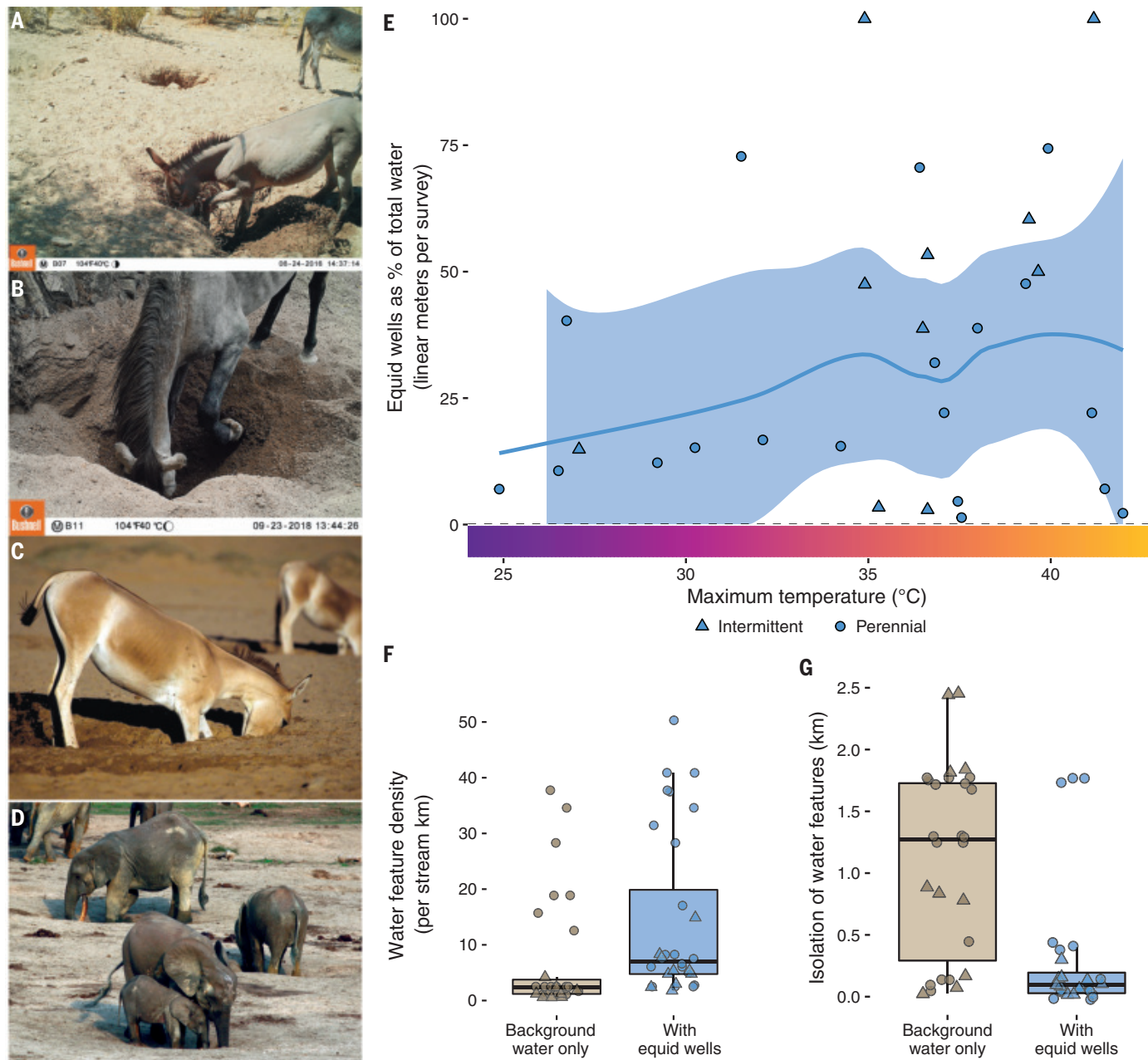
Fig. 1. Well digging increases water availability in desert streams. (**A** to **D**) Well digging by (A) wild donkeys (*E. africanus asinus*) and (B) wild horses (*E. ferus caballus*) in the Sonoran Desert, USA; (C) khulan (*Equus hemionus*) in Mongolia; and (D) African elephants (*Loxodonta africana*). (**E**) The contribution of equid wells increases with temperature. Each point shows the percentage of total water provided by equid wells at each survey and site, measured in meters parallel to the direction of stream flow. Line and fill show loess regression and confidence intervals. Maximum daily temperatures were obtained from PRISM (23). (**F**) Equid wells increase the density of water features (per km). (**G**) Equid wells reduce the isolation of water features, measured as average nearest-neighbor distances between waters. For (F) and (G), boxplots indicate median (central line), interquartile range (bottom and top of box), and minimum and maximum values excluding outliers (lines). [Photo credits: E.J.L. [(A) and (B)], P. Kaczensky (C), and R. Ruggiero (D)]

[1]Centre for Compassionate Conservation, Faculty of Science, University of Technology Sydney, Ultimo, NSW, Australia. [2]Center for Biodiversity Dynamics in a Changing World (BIOCHANGE), Department of Biology, Aarhus University, Aarhus, Denmark. [3]Section for Ecoinformatics and Biodiversity, Department of Biology, Aarhus University, Aarhus C, Denmark. [4]School of Life Sciences, Arizona State University, Tempe, AZ, USA. [5]School of Sustainability, Arizona State University, Tempe, AZ, USA. [6]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA. [7]Roger Williams Park Museum of Natural History, Providence, RI, USA.
*Corresponding author. Email: erick.lundgren@gmail.com
†Deceased.

springs), wells provided up to 74% of surface water by accessing the water table in dry reaches (Fig. 1E). Likewise, equid wells increased water density relative to background water by an average of 332% (SD = 416%) and by as much as 1450% (Fig. 1F).

Isolated water features can be areas of heightened antagonistic interactions among wildlife, including predation, disease transmission, competition, and herbivory (9–11). Equid wells strongly reduced the isolation of water features, reducing average nearest-
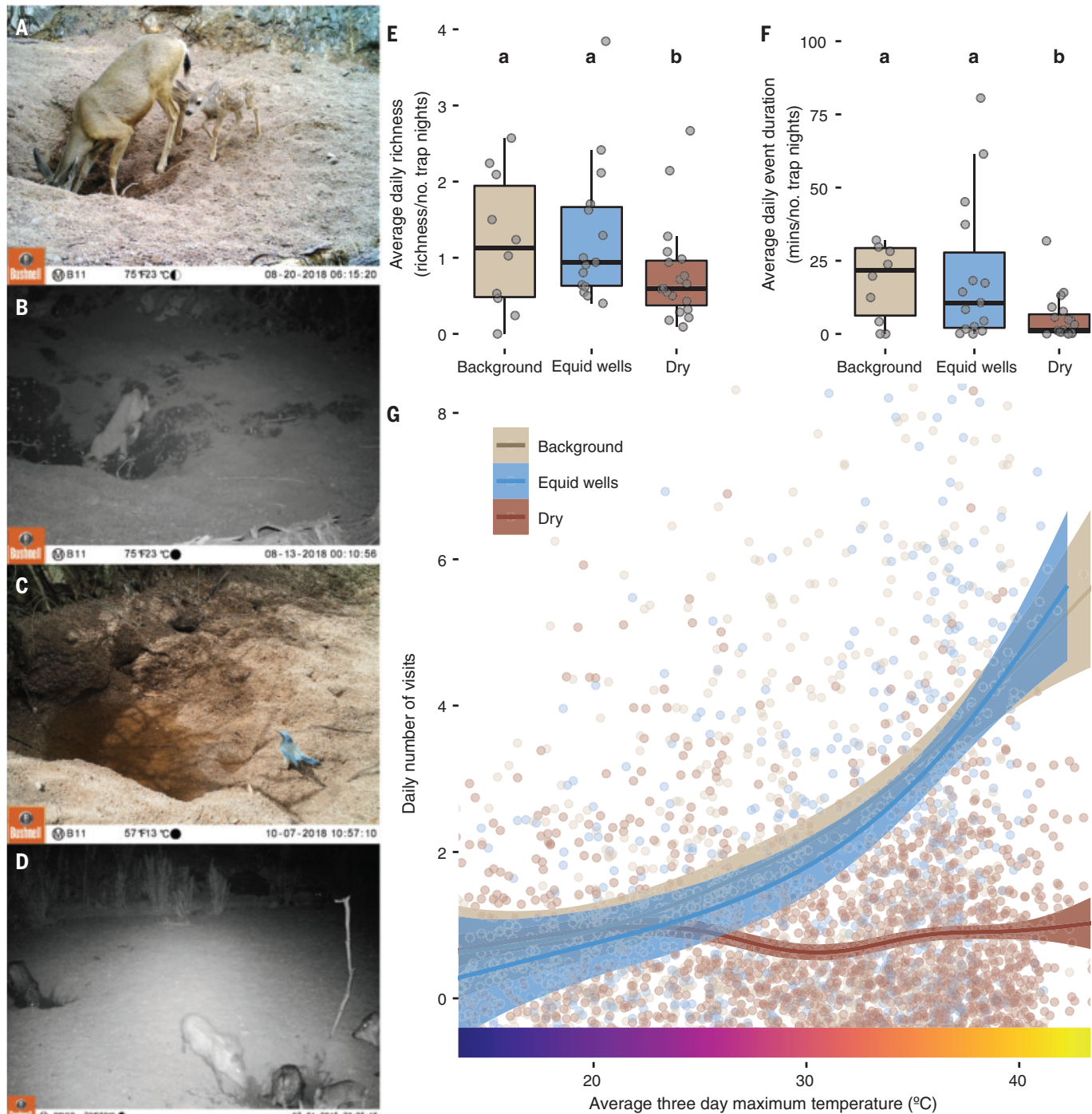
Fig. 2. Use of equid wells by vertebrates. (**A** to **D**) Use of equid wells by select species: (A) mule deer (*O. hemionus*), (B) bobcat (*Lynx rufus*), (C) Woodhouse's scrub-jay (*Aphelocoma woodhouseii*), and (D) javelina (*P. tajacu*). (**E** and **F**) (E) Average daily species richness and (F) daily visit duration, by camera station, relativized by number of trap nights. Letters (a and b) indicate significance groupings (*p* < 0.001). Boxplots indicate median (central line), interquartile range (bottom and top of box), and minimum and maximum values excluding outliers (lines). (**G**) Daily visit frequency by average 3-day maximum temperature (°C). Points indicate days. Line and fill illustrate generalized linear model of relationship between temperature and visit frequency. Equids and vertebrates <100 g (e.g., small passerines and mice) were excluded from analyses, the latter because of poor detection probabilities for small species. Domestic cattle were analyzed separately because of high, but inconsistent, stocking rates throughout the season (fig. S1). [Photo credits: E.J.L. [(A) to (D)]

Fig. 3. Equid wells function as germination nurseries for riparian pioneer trees.
(**A**) Fremont cottonwood (*P. fremontii*) in a series of abandoned equid wells. (**B**) Seedling density between riverbank germination zones, undisturbed surfaces, and equid wells. Filled intervals behind points indicate density distribution of seedlings. Letters (a and b) indicate significance groupings ($p < 0.01$). [Photo credit: E.J.L.]
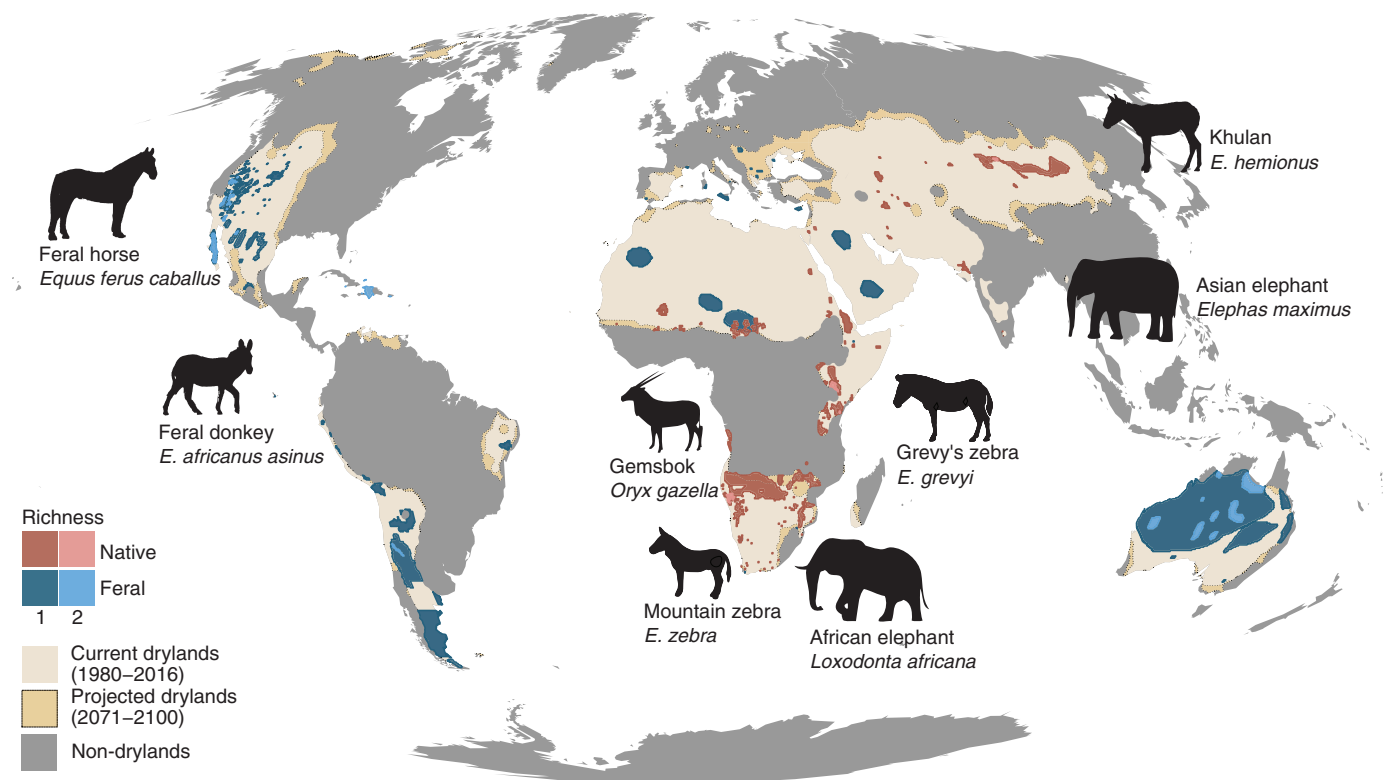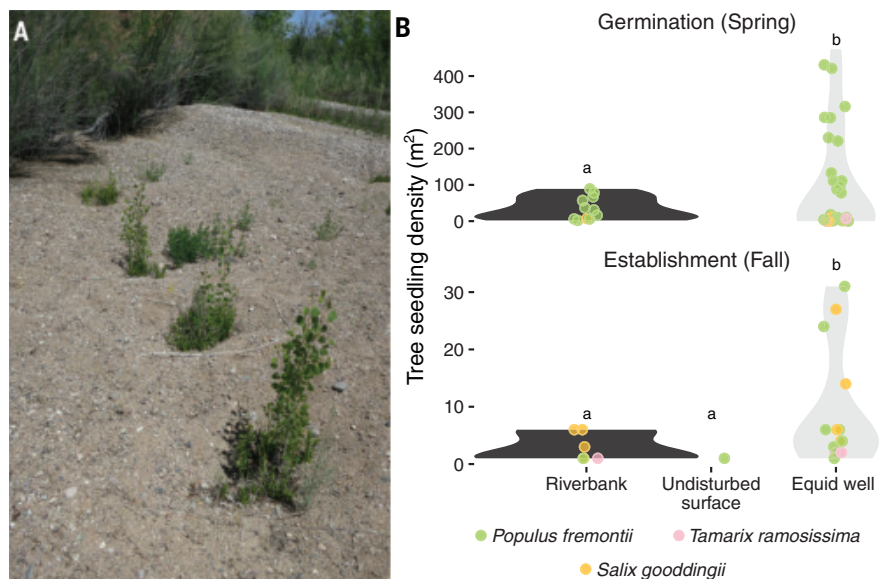
Fig. 4. Well-digging megafauna contribute the capacity to buffer water availability across many drylands. Species richness (number of species) of native and introduced megafauna known to regularly dig wells ≥0.5 m in depth overlaid on current (beige) and projected (yellow ochre) global drylands [semi-arid plus arid Köppen-Geiger climate zones (*19*)]. Gray fill indicates non-drylands. Feral equids (blue) have been documented regularly digging wells in suitable terrain in North America and Australia [table S1; ranges from (*24*, *25*)]. Projected dryland expansion is based on business-as-usual emissions [Representative Concentration Pathway (RCP) 8.5].

neighbor distances between water features by an average of 65% (an 843-m reduction, SD = 798 m) and by as much as 99% (a 2.3-km reduction) (Fig. 1G). Thus, in addition to increasing the total amount of water available (Fig. 1E), we argue that equid wells may relax

the potential for strong antagonistic interactions and reduce the distances that animals must travel to reach water.

To understand whether equid wells have value for other species, we deployed camera traps at five sites in the Sonoran and Mojave

Deserts, sampling over 3258 trap nights (table S2). We calculated daily species richness and the duration and frequency of visits as measures of vertebrate activity, which we compared between equid wells, background waters, and adjacent dry controls. We excluded the

equids themselves and species <100 g to control for poor detection probabilities for small animals.

Overall, we detected 59 vertebrate species at equid wells, of which 57 were recorded drinking [Fig. 2, A to D; table S3; data S2; and (8)]. Daily species richness was 64 and 51% higher on average at equid wells and background waters, respectively, than at dry controls (post hoc test: t ratio = −7.6 and −10.3, $p < 0.0001$; Fig. 2E). Likewise, visit duration was 274 and 620% longer at equid wells and background waters than at dry controls (t ratio = −7.8 and −9.7, $p < 0.0001$; Fig. 2F), and visit frequency was 91 and 60% higher on average (t ratio = −11.3 and −16.32, $p < 0.0001$; Fig. 2G). These effects increased with temperature at both water types (Fig. 2G; see table S4 for omnibus test results). There was no difference between equid wells and background waters in terms of richness (post hoc test: t ratio = −1.6, $p = 0.25$) or duration (t ratio = 1.9, $p = 0.14$), but background waters had higher visit frequencies than equid well waters (t ratio = −3.4, $p = 0.002$).

Responses were similar across vertebrate body sizes (fig. S1), and species composition did not vary between resource types [permutational multivariate analysis of variance (PERMANOVA): $F = 0.9$, $p = 0.7$; fig. S2]. Domestic cattle activity showed a similar pattern but was analyzed separately, because cattle densities were determined by humans and fluctuated during the study period (fig. S3 and table S4).

Some of the species detected drinking from equid wells are free-water dependent, such as mule deer (*Odocoileus hemionus*), whose summer distribution is shaped by surface water [table S3 and (12)]. Other desert species, such as javelina (*Pecari tajacu*), are capable of extracting sufficient water from food and thus can be free-water independent (table S3). However, this requires increased food consumption rates, especially at high temperatures (13). By maintaining water availability, equid wells may both enable the persistence of free-water–dependent species and prevent transitions from "food web" to "water web," thereby reducing the potential for oscillatory population dynamics (14).

Well digging also influences vegetation. On a dammed perennial Sonoran Desert river, abandoned equid wells host numerous riparian trees (primarily *Populus fremontii* and *Salix gooddingii*; Fig. 3A), members of a small-seeded, fast-growing, flood-adapted functional group whose germination requires moist substrate without competing vegetation and whose conservation is considered a regional and global priority (7, 15, 16). We collected data on germination and establishment of these trees across available habitats and found higher seedling density in equid wells than in adjacent

moist riverbank zones (89% of seedlings, Wilcoxon signed rank test: $W = 109$, $p = 0.007$; Fig. 3B), a difference which persisted after typical summer mortalities [generalized linear mixed-effect model: $\chi^2 = 54.6$, $p < 0.0001$; Fig. 3B, data S2, and (8)].

Although riverbanks provided the moist substrate necessary for germination, they were significantly more herbaceous ($p < 0.0001$; fig. S4A), and seedling density showed a strong negative relationship to herbaceous cover ($p < 0.0001$; fig. S4B). This suggests that riverbanks are less competitively suitable for these small-seeded pioneer trees, a relationship shown in previous work (7, 16). Thus, by exposing moist substrate free of competing herbaceous vegetation, we argue that equid wells can serve as flood-mimicking nurseries. The importance of megafauna for the dispersal of large-seeded, high–wood density functional groups is well known (4), yet facilitation of pioneer trees through megafauna disturbance is less recognized. Further research is necessary to understand how these nurseries may contribute to long-term riparian forest dynamics.

Equid well digging was limited by water-table depth, with equids unlikely to dig deeper than 2 m (table S1). Well digging was also constrained by substrate, primarily occurring in flood-disturbed systems of loose sand and gravel ($p < 0.0001$; fig. S5). The ecological relevance of this behavior also appears to be shaped by the availability of alternative water sources in the broader landscape and by stream intermittency. Intermittent streams, the most common stream type across nearly half of Earth (17), were where equid wells had the strongest effects on water availability. Stream intermittency is projected to increase as currently perennial streams lose yearlong flows (18) and as drylands expand [Fig. 4 and (19)] as result of groundwater mining, agriculture, and climate change (17). These reductions in water availability, coupled with rising temperatures, are projected to have strong effects on biodiversity and ecosystem function (6). Our results suggest that equids and other well-digging megafauna have the potential to mitigate these changes, whether native or introduced (Fig. 4).

Recent and ancient extinctions and range contractions of megafauna, and the loss of their distinct ecological functions, has led to highly modified modern landscapes (4). Although introduced megafauna have primarily been studied as threats to conservation goals (20), growing evidence suggests that they present a countercurrent to ancient losses (21) and may replace lost ecological functions (22). Specifically, we show that feral equids can increase water availability in drylands, with associated effects on a variety of species and ecosystem processes. We suggest that well digging by feral equids may replace a func-

tion lost with the extinctions of large vertebrates across the world's drylands.

REFERENCES AND NOTES

1. W. J. Ripple et al., *Sci. Adv.* **1**, e1400103 (2015).
2. A. D. Barnosky, P. L. Koch, R. S. Feranec, S. L. Wing, A. B. Shabel, *Science* **306**, 70–75 (2004).
3. J. L. Metcalf et al., *Sci. Adv.* **2**, e1501682 (2016).
4. Y. Malhi et al., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 838–846 (2016).
5. W. H. Schlesinger et al., *Science* **247**, 1043–1048 (1990).
6. K. E. McCluney et al., *Biol. Rev. Camb. Philos. Soc.* **87**, 563–582 (2012).
7. J. Stromberg, V. Beauchamp, M. Dixon, S. Lite, C. Paradzick, *Freshw. Biol.* **52**, 651–679 (2007).
8. E. J. Lundgren et al., Equids engineer desert water availability. Figshare (2021); https://doi.org/10.6084/m9.figshare.13668794.
9. T. C. Atwood, T. L. Fry, B. R. Leland, *J. Wildl. Manage.* **75**, 1609–1615 (2011).
10. J. D. Rogerson, W. S. Fairbanks, L. Cornicelli, *J. Wildl. Dis.* **44**, 28–44 (2008).
11. I. Thrash, G. K. Theron, J. P. Bothma, *J. Arid Environ.* **29**, 213–219 (1995).
12. K. R. Rautenstrauch, P. R. Krausman, *J. Mammal.* **70**, 197–201 (1989).
13. D. C. Allen, K. E. McCluney, S. R. Elser, J. L. Sabo, *Front. Ecol. Environ.* **12**, 156–160 (2014).
14. K. E. McCluney, J. L. Sabo, *Ecology* **90**, 1463–1469 (2009).
15. E. González et al., *J. Environ. Manage.* **218**, 374–387 (2018).
16. P. B. Shafroth et al., *Ecol. Eng.* **106**, 645–660 (2017).
17. T. Datry, B. Núria, B. Andrew, Eds., *Intermittent Rivers and Ephemeral Streams: Ecology and Management* (Elsevier, 2017).
18. A. L. Burrell, J. P. Evans, M. G. De Kauwe, *Nat. Commun.* **11**, 3853 (2020).
19. H. E. Beck et al., *Sci. Data* **5**, 180214 (2018).
20. A. D. Wallach, E. J. Lundgren, W. J. Ripple, D. Ramp, *Conserv. Biol.* **32**, 962–965 (2018).
21. E. J. Lundgren, D. Ramp, W. J. Ripple, A. D. Wallach, *Ecography* **41**, 857–866 (2018).
22. E. J. Lundgren et al., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 7871–7878 (2020).
23. PRISM Climate Group, Oregon State University; http://prism.oregonstate.edu.
24. A. D. Wallach et al., *Conserv. Biol.* **34**, 997–1007 (2020).
25. International Union for the Conservation of Nature (IUCN), IUCN Red List of Threatened Species (IUCN, 2018).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6541/491/suppl/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S4
References (26–74)
MDAR Reproducibility Checklist
Data S1 and S2

MULTIFERROICS

Electric field control of natural optical activity in a multiferroic helimagnet

Ryoji Masuda[1], Yoshio Kaneko[2], Yoshinori Tokura[1,2,3], Youtarou Takahashi[1,2]*

Controlling the chiral degree of freedom in matter has long been an important issue for many fields of science. The spin-spiral order, which exhibits a strong magnetoelectric coupling, gives rise to chirality irrespective of the atomic arrangement of matter. Here, we report the resonantly enhanced natural optical activity on the electrically active magnetic excitation, that is, electromagnon, in multiferroic cupric oxide. The electric field control of the natural optical activity is demonstrated through magnetically induced chirality endowed with magnetoelectric coupling. These optical properties inherent to multiferroics may lead to optical devices based on the control of chirality.

Chirality is a fundamental concept in many fields, including physics, chemistry, and biology. The asymmetry of spatial coordinates of constituent atoms is responsible for the emergence of chirality in molecules, crystalline solids, and metamaterials (1–3). In solids, additional mechanisms are possible, such as spin-spiral long-range magnetic orders, which can break the space-inversion symmetry irrespective of the underlying crystalline symmetry. Spin-spiral orders are gaining importance in the context of enhanced magnetoelectric coupling in multiferroics and of magnetic skyrmions, which are of interest for various applications (4–8). Chiral magnetic orders in centrosymmetric atomic lattices have a twofold degeneracy, i.e., left- and right-handed chirality, which can be controlled through magnetic phase transitions or by driving chiral domain walls in external fields, in contrast to the rigid handedness of chiral atomic structures.

Figure 1, A and B, shows the schematics of two spin-spiral orders, proper screw spin and cycloidal spin structures, respectively. The screw spin structure with the spin plane perpendicular to the screw axis (Fig. 1A) exhibits chiral asymmetry: The mirror image of the left-handed screw is the right-handed one, with opposite chirality. The essential difference from the usual crystallographic chirality is that the helix of the spins with axial vector character, not the asymmetry of spatial coordinates of spins, produces the chirality. The cycloidal spin structure in Fig. 1B also breaks the space inversion symmetry but the chirality is absent; the mirror operation and the 180° rotation along screw axis are identical for this spin structure while shifting the spin rotational phase. The electric polarity induced by this cycloidal spin order gives rise to a magnetically induced ferroelectricity and a strong static and/or dynamical magnetoelectric coupling, as demonstrated in many multiferroics (4–6). Although the existence of the chiral degree of freedom for the screw spin order (Fig. 1A) has been suggested in terms of the spin structure (9, 10), it remains an open question whether a magnetically induced chiral asymmetry would exhibit enhanced electronic and optical properties.

Symmetry breaking can give rise to previously unobserved optical effects, which are resonantly enhanced upon the creation of relevant excitations. For example, directional dichroism and optical rotation upon spin excitations have been reported for several multiferroics with a cycloidal spin order (11–14). The magnetic field–induced optical activity is also demonstrated in multiferroics with special crystalline symmetries such as the D_{2d} point group (15, 16). Upon the creation of spin excitations, the spontaneous breaking of chiral symmetry arising from the long-range screw spin order potentially produces the enhanced optical effect inherent to chiral asymmetry, i.e., natural optical activity (NOA) (1, 2). The observation of this phenomenon, which is to be distinguished from the NOA caused by magnetic excitations in chiral atomic lattices (17), remains elusive.

Here, we focused on the optical property associated with a magnetically induced chirality in a multiferroic helimagnet, in which the different orders, ferroelectricity and chirality, are strongly coupled. We studied the spin-spiral order of the high-temperature multiferroic CuO (18) and investigated its NOA on the electrically active magnetic resonance, i.e., electromagnon resonance, arising from the spin-spiral order in the terahertz region (19, 20). The emergence of enhanced NOA was identified on the enantiopure chiral domain formed by the electric field cooling. We also demonstrated the temporal control of the polarization of light through NOA by driving the chiral domain wall in isothermal condition.

Between 213 and 230 K, the multiferroic CuO exhibits spin-spiral order with a canted spin plane, which is a superposition of the screw and cycloidal spin structures, as shown in Fig. 1C; a collinear antiferromagnetic (AF1) phase emerges at <213 K (Fig. 1D) (9, 21–24). The spin-spiral order breaks the space inversion symmetry in the centrosymmetric C2/c monoclinic crystalline lattice. The cycloidal spin component induces the ferroelectric polarization $\boldsymbol{P}$ along the [010] axis (Fig. 1E), which is expressed as $\boldsymbol{P} \propto \boldsymbol{e}_{ij} \times (\boldsymbol{S}_i \times \boldsymbol{S}_j)$, where $\boldsymbol{e}_{ij}$ is a unit vector connecting the neighboring spins $\boldsymbol{S}_i$ and $\boldsymbol{S}_j$ (25–27). On the other hand, the screw spin component induces the chirality. In CuO, the angle between the spin plane and the magnetic modulation vector $\boldsymbol{q}_m$ is 107° (Fig. 1E), indicating that the screw spin component is more dominant than the cycloidal one (9). It should be noted that the cross-product of spins, $\boldsymbol{S}_i \times \boldsymbol{S}_j$, is sometimes referred to as the vector spin chirality; however, it does not necessarily induce chirality. Because the helicity of the spin-spiral order is responsible for the sign of $\boldsymbol{P}$, the magnetically induced chirality is rigidly connected to $\boldsymbol{P}$, enabling the electrical control of the chirality, as suggested by the neutron-scattering measurements (9, 10). This property should be distinguished from the spin-spiral order in chiral crystalline lattices, in which the helicity of the spin order is uniquely determined by the chirality of the crystal.

To address the optical response caused by the magnetically induced chirality, we focused on the electromagnon resonance in the terahertz region. Because the fluctuations of the spin-spiral order produce a transition electric dipole, the magnon resonance exhibits electric activity. In fact, the dynamical magnetoelectric coupling inherent to the electromagnon resonance produces the gigantic nonreciprocal directional dichroism (11–13, 28). The NOA is observed as the optical rotation for light propagating in chiral media. Therefore, we used time-domain terahertz polarimetry to demonstrate NOA. The temporal profiles of terahertz pulses, including amplitude and phase, were measured with the wire grid polarizer 2 (Fig. 2A), which was set to be parallel ($E^\omega_{//}$) or perpendicular (crossed-Nicol geometry: E^ω_{CN}) to the polarization of the incident light (E^ω_i). These polarization measurements were done for the incident light with $E^\omega_i //$[101] and $E^\omega_i //$[010], which are eigen polarizations in the paramagnetic phase. To label the optical constants, we used the experimental coordinates x and y for [101] and [010], respectively. The obtained datasets provide the 2×2 complex dielectric spectra including the diagonal [$\varepsilon_x(\omega)$, $\varepsilon_y(\omega)$] and the off-diagonal [$\varepsilon_{xy}(\omega)$, $\varepsilon_{yx}(\omega)$] terms. In this analysis, we omitted the tilt of the wave front for extraordinary ray ($E^\omega_e //$ [101]), and the wave front was assumed to be parallel to the sample surface.

[1]Department of Applied Physics and Quantum Phase Electronics Center (QPEC), University of Tokyo, Tokyo 113-8656, Japan. [2]RIKEN Center for Emergent Matter Science (CEMS), Wako, Saitama 351-0198, Japan. [3]Tokyo College, University of Tokyo, Tokyo 113-8656, Japan.
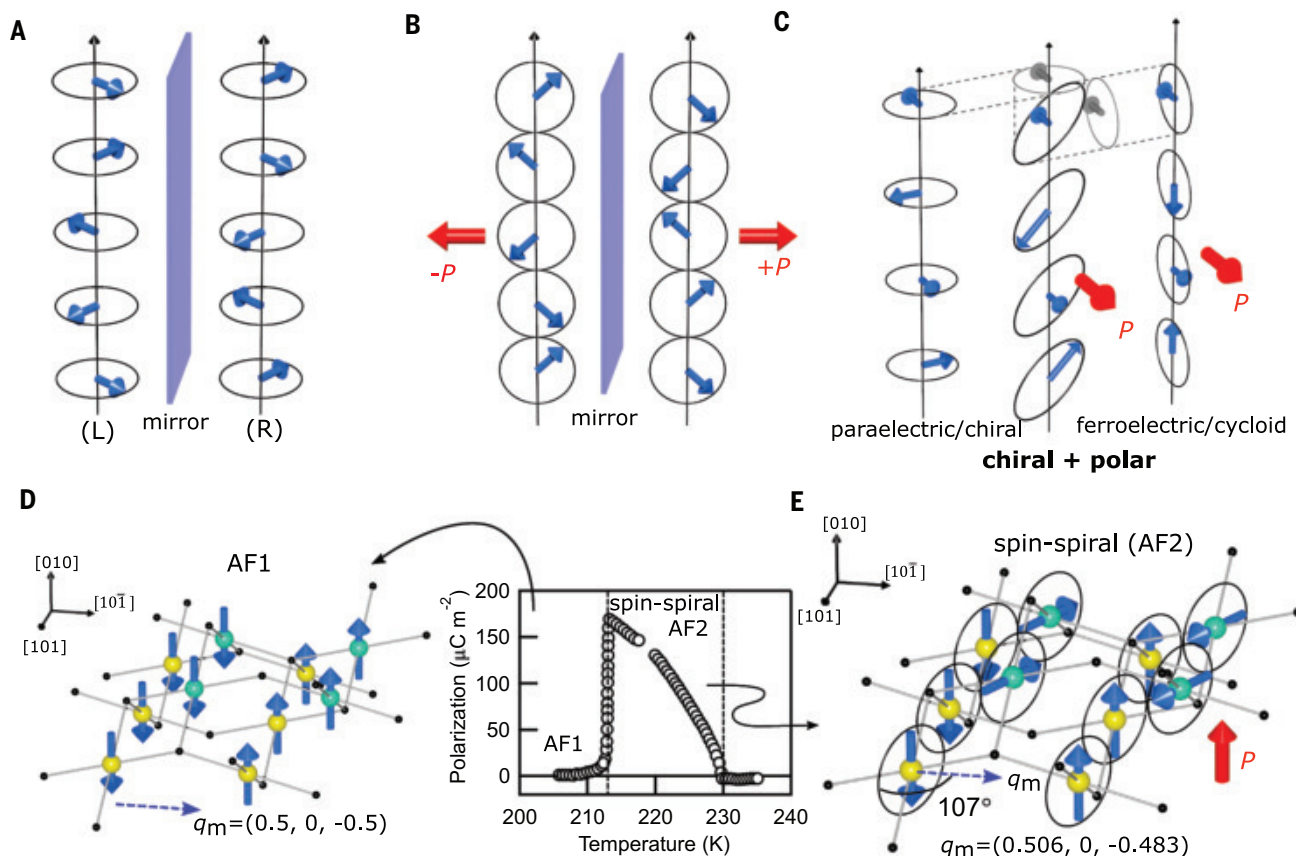*Corresponding author. E-mail: youtarou-takahashi@ap.t.u-tokyo.ac.jp

Fig. 1. Chirality arising from spin-spiral structure in multiferroic CuO.
(**A**) Screw type of spin spiral. The left-handed spin helix (L) has opposite chirality
to its mirror image, the right-handed spin helix (R). (**B**) Cycloidal spin structure.
Its polar nature leads to the spontaneous polarization ±*P*; chirality is absent.
(**C**) Superposition of the screw and cycloidal spin structures is a spin spiral with
canted spin plane that has both chirality and polarity, as realized in the
spin-spiral (AF2) phase in CuO. (**D**) In the AF1 phase of CuO (<213 K), the
collinear up-up-down-down spin structure emerges with a magnetic modulation

vector q_m = (0.5, 0, −0.5), as indicated by the dashed blue arrow. To clearly
indicate the equal phase plane of spin orders, the Cu sites are colored yellow
and light blue. The temperature dependence of the ferroelectric polarization
is reproduced from (*18*). (**E**) Spin-spiral phase with a canted spin plane emerges
at intermediate temperatures (between 213 and 230 K) with an incommensurate
magnetic modulation vector q_m = (0.506, 0, −0.483). The angle between q_m
and spin plane is 107°. The cycloidal spin component is responsible for the
ferroelectric polarization *P*.

The optical spectra in the spin-spiral phase
(213 to 230 K) showed a pronounced peak
structure around 3 meV for $\varepsilon_x(\omega)$ (//[101]) (Fig.
2B), whereas no clear resonance was observed
for the paramagnetic (>230 K) and the AF1
phases (<213 K) as previously reported (*29*).
This resonance in the spin-spiral phase has
been ascribed to the electromagnon, the
electric activity of which is attributed to the
fluctuations of magnetically induced *P* (*30*).
For the spectra $\varepsilon_y(\omega)$ (//[010]), the resonance
structure was absent, including in the spin-
spiral phase. To realize a single chirality domain
state, the electric field was applied along [010]
while cooling the sample [see (*31*) for uniformity
of electric field]. Thus, we were able to examine
the optical spectra in the single chirality domain
without applying an external electric or mag-
netic field. In the spin-spiral phase, optical
rotation was found for the incident light with
E_i^ω//x, as shown in the complex rotation spectra
$\theta(\omega) + i\eta(\omega)$ (Fig. 2C); time-domain terahertz
polarimetry directly provides the rotation angle

$\theta(\omega)$ and ellipticity $\eta(\omega)$ (*31*). The rotation angle
$\theta(\omega)$ amounts to 35 mrad (~2°) for the sample
with thickness of 1.1 mm at 214 K. Optical
rotation was also observed for E_i^ω//y (Fig. 2D),
although this was in the off-resonant geome-
try [see $\varepsilon_y(\omega)$ in Fig. 2B]. The rotation angles
were enhanced on the resonance of the electro-
magnon around 3 meV in both geometries,
but the difference in the spectral shapes was
retained. This difference was caused by crys-
tallographic birefringence. By contrast, no
polarization rotation was observed in the low-
temperature AF1 phase (212 K) or in the para-
magnetic (PM) phase (231 K), as expected.

To examine the NOA in an anisotropic me-
dium, we deduce the 2 × 2 effective dielectric
tensor spectra $\tilde\varepsilon(\omega)$ [see fig. S3 for the spectra
for $\varepsilon_{xy}(\omega)$ and $\varepsilon_{yx}(\omega)$]. The NOA causes cir-
cular birefringence/dichroism with k^ω depend-
ence for light propagating in chiral media,
indicating that the NOA is different from
genuine dielectric phenomena. As we discuss
later, the NOA originates from the dynamical

magnetoelectric coupling. The extraordinary
ray in the monoclinic crystal possibly affects
the optical rotation especially for the reso-
nance. Nevertheless, this formalism provides
evidence for the emergence of the circular
birefringence/dichroism as well as a quantita-
tive evaluation of the optical rotatory power
for the observed NOA. The magnetic and mag-
netoelectric activities are implicitly included
in the effective dielectric spectra. The optical
rotation for light propagating in chiral media
can be expressed by the antisymmetric part
of the effective dielectric tensor $i\varepsilon_A$, which is
purely imaginary in a dissipation-less medium,
whereas the symmetric term ε_S represents the
linear birefringence as follows (*32*):

$$\tilde\varepsilon(\omega) = \begin{pmatrix} \varepsilon_{xx}(\omega) & \varepsilon_{xy}(\omega) \\ \varepsilon_{yx}(\omega) & \varepsilon_{yy}(\omega) \end{pmatrix}$$
$$= \begin{pmatrix} \varepsilon_x(\omega) & i\varepsilon_A(\omega) + \varepsilon_S(\omega) \\ -i\varepsilon_A(\omega) + \varepsilon_S(\omega) & \varepsilon_y(\omega) \end{pmatrix}$$

The resultant complex spectra of $\varepsilon_A(\omega)$, which
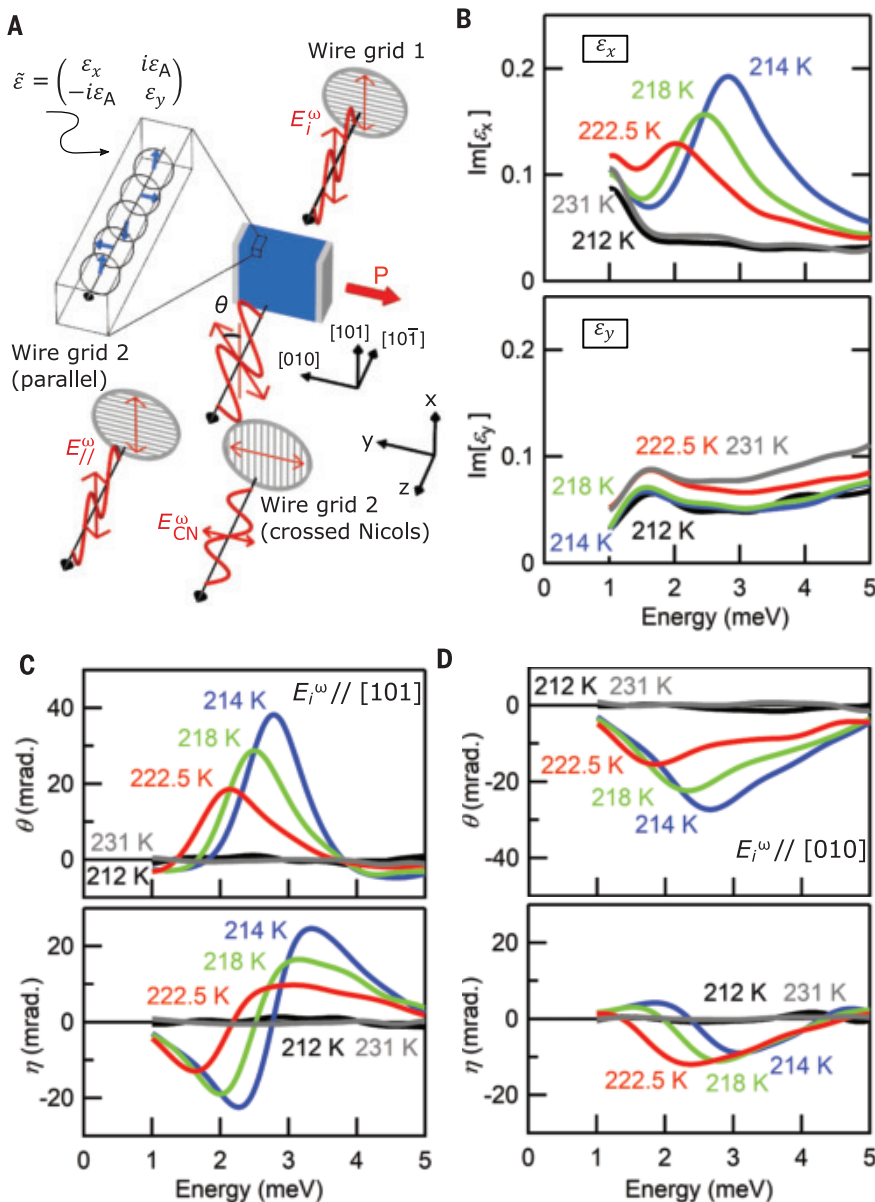are obtained by antisymmetrizing the observed

Fig. 2. Observation of enhanced optical rotation on electromagnon resonance. (**A**) Experimental setup for the terahertz time-domain polarimetry measurements. For the full polarization analysis, the phase and amplitude of the terahertz wave were measured with the crossed-Nicol (E_{CN}^{ω}) and parallel ($E_{//}^{\omega}$) configurations. (**B**) Electromagnon resonance shows up in the spin-spiral phase for ε_x(||[101]) with a pronounced peak at 2.8 meV (214 K), whereas the spectra for ε_y show no resonance structure and little temperature dependence. (**C** and **D**) The complex rotation spectra including the rotation angle θ and ellipticity η for the incident light polarized along [101] (C) and [010] (D). The optical rotation is observed on the electromagnon resonance in the spin-spiral phase, whereas the AF1 (212 K) and PM (231 K) phases show no optical rotation.

off-diagonal components as $i\varepsilon_A(\omega) = [\varepsilon_{xy}(\omega) - \varepsilon_{yx}(\omega)]/2$, are shown in Fig. 3, A and B. The Re $[\varepsilon_A(\omega)]$ and Im$[\varepsilon_A(\omega)]$ are responsible for the circular birefringence and circular dichroism, respectively. The peak (Fig. 3B) and dispersive (Fig. 3A) structures indicate that the resonance with optical activity emerges in the spin-spiral phase. By contrast, no optical activity was discerned in the achiral PM (231 K) or AF1 (212 K) phases. Therefore, these results indicate that the magnetically induced chirality gives rise to the NOA, leading to the optical rotation and the resonance structure in $\varepsilon_A(\omega)$. The peak intensity cf. Im$[\varepsilon_A(\omega)]$ at 214 K (Fig. 3B) was as large as 0.02, which amounts to 10% of peak magnitude of electromagnon in $\varepsilon_x(\omega)$ (Fig. 2B). Conversely, the symmetric

part $\varepsilon_S(\omega) = [\varepsilon_{xy}(\omega) + \varepsilon_{yx}(\omega)]/2$ shows neither energy nor temperature dependence (fig. S4), as expected. Thus, the possible electro-optic effect or magnetostriction, which can induce optical rotation because of linear birefringence, are excluded. Here, the optical rotation arising from the magnetic optical activity, i.e., Faraday effect, can be also excluded, because the time-reversal symmetry is preserved for the spin spiral with no net magnetization.

The color-coded spectra in Fig. 3, C to F, show the correlations among the electromagnon, NOA, and the magnetic phase diagram. The spin-spiral order produces the ferroelectric polarization P (Fig. 3C) and the electromagnon resonance (Im$[\varepsilon_x]$ in Fig. 3D). The NOA (Re$[\varepsilon_A]$ and Im$[\varepsilon_A]$ in Fig. 3, E and F, respectively) has

the same features as the electromagnon: a peak shift toward lower energy and a reduction of the intensity with increasing temperature in the spin-spiral phase. The NOA, electromagnon, and P are all absent in the AF1 and PM phases. These results confirm the essential role of the spin-spiral order for the NOA on the electromagnon. The small peaks at 1 meV in $\varepsilon_x(\omega)$ (Figs. 2B and 3D) and at 1.5 meV in $\varepsilon_y(\omega)$ (Fig. 2B) have no correlation with the magnetic phase transitions, and their origin is not clear.

In general, the magnitude of NOA for the transition from the ground state (g) to the excited state (e) is expressed by the product of electric transition dipole μ and magnetic transition dipole m as $-i\langle g|\mu|e\rangle\langle e|m|g\rangle$
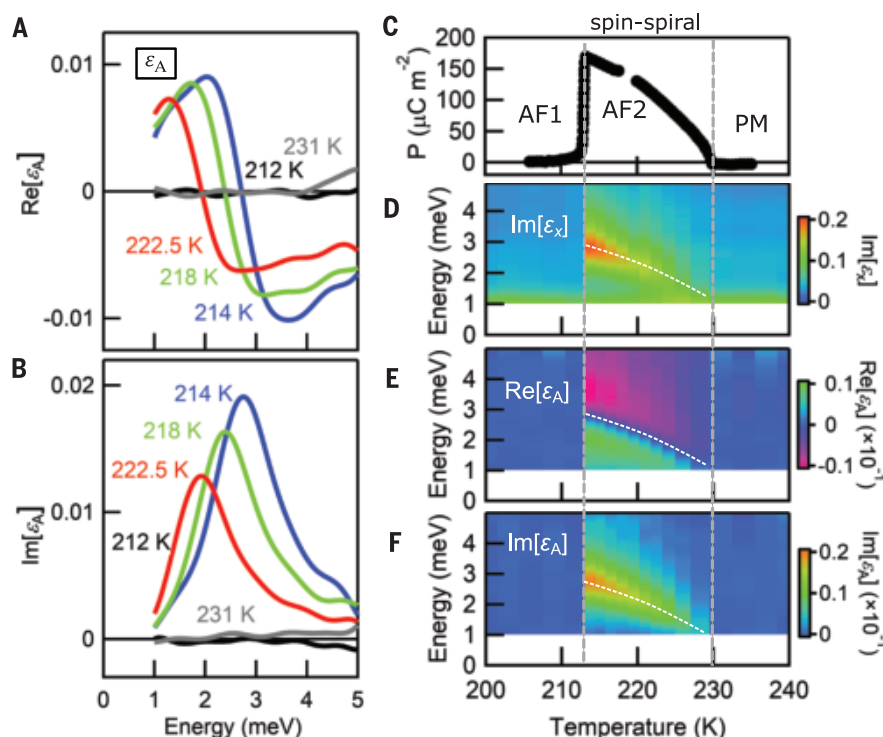
Fig. 3. NOA arising from magnetically induced chirality. The antisymmetric part of the effective dielectric tensor composed of Re[ε_A] (**A**) and Im[ε_A] (**B**). The clear resonance structure with circular birefringence and circular dichroism manifests the resonantly enhanced NOA. Shown is the correlation between phase diagram with the ferroelectric polarization **P** (**C**), electromagnon resonance (Im[ε_x]) (**D**), and the NOA, Re[ε_A] (**E**) and Im[ε_A] (**F**). The spin-spiral order produces the **P**, electromagnon, and NOA. The temperature dependence of the spectral feature of NOA [(E) and (F)] is consistent with that of the electromagnon (D), as indicated by the shift of center frequency (dashed lines).

(*2*, *33*). The resonantly enhanced NOA (Fig. 3, A and B) indicates that strong interference between **m** and **μ** emerges on the electromagnon resonance. In spin-spiral orders, including the genuine screw spin order, the magnon with the fluctuation of spin plane always accompanies the electric transition dipole **μ** similar to the spin-driven ferroelectricity (*21*, *25–28*), in addition to the intrinsic magnetic transition dipole **m**. In many multiferroics, it is reported that enhanced dynamical magnetoelectric coupling produces nonreciprocal directional dichroism (*11–13*, *15*, *16*, *28*, *34*), which can be regarded as the dynamical counterpart of the linear magnetoelectric coupling. By contrast, the NOA is a genuine dynamical effect that disappears in the dc limit and is thus distinguished from the optical magnetoelectric effects. The resonance of the toroidal moment may have some contribution to the NOA; in fact, the antiferroic order of the toroidal moment has been reported in the AF1 phase of the CuO (*35*, *36*).

Electric control of the light polarization was demonstrated in the isothermal condition through the manipulation of the magnetically induced chirality. Figure 4A shows the rotation angle of the light polarization at 2.5 meV versus the static electric field along [010] at 218 K [see (*31*) for the evaluation of the rotation angle in Fig. 4]. In the spin-spiral phase, the crystals prepared without an electric field are in a mixed domain state in terms of

chirality, i.e., racemic mixture, because of the degeneracy of the right- and left-handed spin-spiral orders. The equal amounts of the right- and left-handed chiral domains cancel the rotation of the light polarization as shown in (i) in Fig. 4A (*37*). Because the electric field drives the chiral (and polar) domain wall, the domain growth and coarsening occurs with increasing the electric field from 0 to +50 V/mm, leading to the rise of the rotation angle as indicated by (ii) in Fig. 4A. Above +50 V/mm, the single domain state, i.e., enantiopure spin-spiral order, is produced, as indicated by the saturation of the rotation angle [(iii) in Fig. 4A]. The enantiopure magnetic state is non-volatile, i.e., stable in zero electric field [(iv) in Fig. 4A]. The negative electric field drives the chiral domain wall, and the switching of the chirality is manifested by the reversal of rotation angle below –100 V/mm [(v) in Fig. 4A]. The coercive field for the switching of the chiral and polar domain is ~50 V/mm, which is consistent with the reported **P**-**E** hysteresis curve (*18*). The resultant hysteresis curve of NOA demonstrates the ferroic nature of magnetically induced chirality. It was reported that the perfect single domain could not be formed by the electric field in CuO (*9*). We observed that the postannealing of the crystal improves the saturation level of the rotation angle with high reproducibility (*31*), although the minority domain might remain above the saturation field.

This flexible chirality provides control of the light polarization through the NOA by applying electric field pulses (Fig. 4B). The magnitude of the electric field pulses was set to 100 V/mm, which produces the single chirality domain out of the spin-spiral order (Fig. 4A). A positive pulse produces persistent enantiopure spin-spiral order, so the transmitted terahertz light exhibits counterclockwise light polarization rotation. By contrast, a negative pulse induces the clockwise light polarization rotation because of the sign reversal of chirality.

We should mention here the possible NOA induced by the electric polarization. The crystal symmetry of the CuO in paramagnetic phase is C2/c, which has the twofold rotation axis parallel to [010] axis (//**P**) and the mirror plane perpendicular to the twofold axis. The electric polarization along [010] removes the mirror symmetry, so that the electric polarization–induced chirality coexists with the chirality stemming from the screw spin structure. In the spin-spiral phase with a canted spin plane in CuO, the roles of the screw and cycloidal spin structures for the NOA cannot be distinguished. We note that the optical rotation arising from the field-induced polarization remains small. The electric field (100 V/mm) induces an additional polarization of 16 μC/m² , which is as large as 10% of the spontaneous polarization at 218 K (150 μC/m²) (*31*). As shown in Fig. 4B and fig. S5, however,
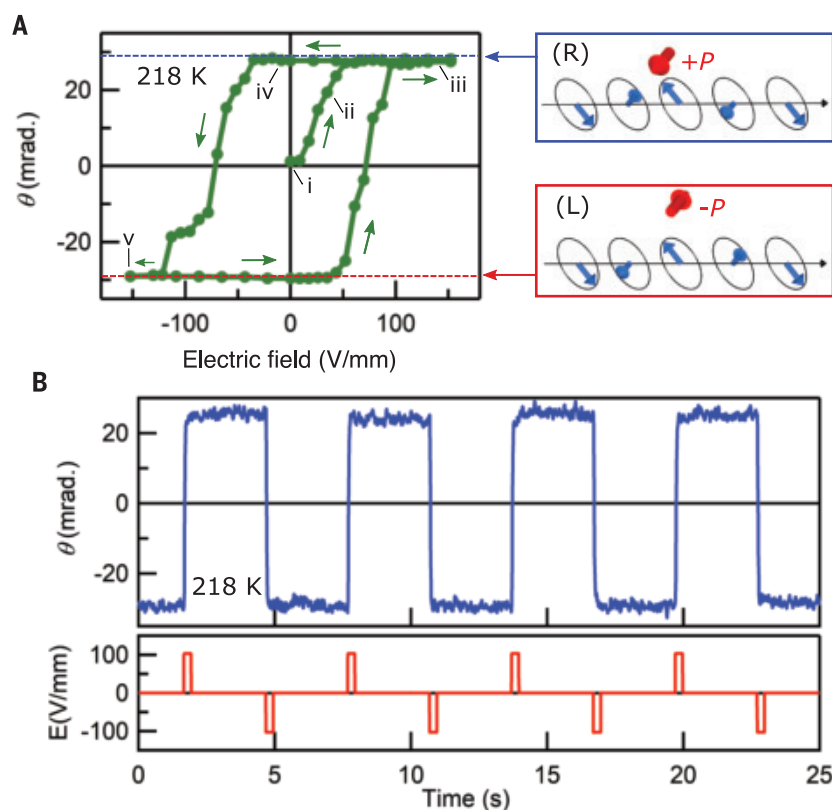
Fig. 4. Electric field control of NOA. (**A**) The NOA or, equivalently, the magnetically induced chirality is controlled by the electric field in isothermal condition (218 K). The rotation angle θ at 2.5 meV (Fig. 2C) is plotted. The hysteretic behavior of the NOA after the zero-field cooling demonstrates the ferrochiral nature of spin-spiral order. The blue and red dashed lines indicate the level of rotation angles obtained by the electric field cooling, i.e., that of the enantiopure domain. The sequence of manipulation of chiral domain is denoted by (i) to (v). (**B**) Temporal control of the NOA (upper panel) by application of the electric field pulses with the height of ±100 V/mm and the pulse duration of 300 ms (bottom panel). The persistent enantiopure spin-spiral order is formed by the electric field pulses, and the switching of the NOA and chirality is demonstrated.

the field-induced dielectric polarization does not enhance the rotation angle within the precision of this measurement (~1 mrad).

A different type of optical rotation has been reported with remanent signal in the electric field sweep in the (achiral) polar cycloidal helimagnet $DyMnO_3$ (*14*). This optical rotation is prohibited for the light propagating parallel to the magnetic modulation vector ($k^\omega // q_m$). In our experiment, k^ω was parallel to [10$\bar{1}$], which was almost parallel to $q_m = [0.506, 0, -0.483]$, so this polar-type optical rotation contributes little to the observed NOA.

The electro-optic functionality described here arises from the breaking of space-inversion symmetry in spin-spiral order, so our results apply to multiferroic helimagnets in general. The chirality switching was accompanied by only a tiny displacement current, requiring low power consumption, in contrast to the light polarization control through magneto-optical (Faraday) effect or the magnetic field–induced NOA (*15*, *16*). This flexible chirality and concomitant NOA inherent to the helimagnet enables multiferroics-based chiral optics and may lead to devices with electric writing and optical reading. The coherent control of magnetically induced chirality in picoseconds could be achieved by excitation with the resonant frequency of electromagnon (*38*), which may lead to optical control of the chiral degree of freedom.

REFERENCES AND NOTES

1. L. M. Pasteur, *Ann. Chim. Phys. Sér.* **24**, 442–459 (1848).
2. G. H. Wagnière, *On Chirality and the Universal Asymmetry* (Wiley, 2007).
3. J. K. Gansel *et al.*, *Science* **325**, 1513–1515 (2009).
4. T. Kimura *et al.*, *Nature* **426**, 55–58 (2003).
5. Y. Tokura, S. Seki, N. Nagaosa, *Rep. Prog. Phys.* **77**, 076501 (2014).
6. F. Matsukura, Y. Tokura, H. Ohno, *Nat. Nanotechnol.* **10**, 209–220 (2015).
7. N. Nagaosa, Y. Tokura, *Nat. Nanotechnol.* **8**, 899–911 (2013).
8. A. Fert, N. Reyren, V. Cros, *Nat. Rev. Mater.* **2**, 17031 (2017).
9. P. Babkevich *et al.*, *Phys. Rev. B Condens. Matter Mater. Phys.* **85**, 134428 (2012).
10. J. Stein *et al.*, *Phys. Rev. Lett.* **119**, 177201 (2017).
11. I. Kezsmarki *et al.*, *Phys. Rev. Lett.* **106**, 057403 (2011).
12. Y. Takahashi, R. Shimano, Y. Kaneko, H. Murakawa, Y. Tokura, *Nat. Phys.* **8**, 121–125 (2012).
13. Y. Takahashi, Y. Yamasaki, Y. Tokura, *Phys. Rev. Lett.* **111**, 037204 (2013).
14. A. Shuvaev *et al.*, *Phys. Rev. Lett.* **111**, 227201 (2013).
15. M. Saito, K. Ishikawa, K. Taniguchi, T. Arima, *Phys. Rev. Lett.* **101**, 117402 (2008).
16. S. Bordács *et al.*, *Nat. Phys.* **8**, 734–738 (2012).
17. A. M. Kuzmenko *et al.*, *Phys. Rev. B Condens. Matter Mater. Phys.* **89**, 174407 (2014).
18. T. Kimura, Y. Sekio, H. Nakamura, T. Siegrist, A. P. Ramirez, *Nat. Mater.* **7**, 291–294 (2008).
19. A. Pimenov *et al.*, *Nat. Phys.* **2**, 97–100 (2006).
20. H. Katsura, A. V. Balatsky, N. Nagaosa, *Phys. Rev. Lett.* **98**, 027203 (2007).
21. S. Åsbrink, L.-J. Norrby, *Acta Crystallogr. B* **26**, 8–15 (1970).
22. B. X. Yang, T. R. Thurston, J. M. Tranquada, G. Shirane, *Phys. Rev. B Condens. Matter* **39**, 4343–4349 (1989).
23. P. J. Brown, T. Chattopadhyay, J. B. Forsyth, V. Nunez, *J. Phys. Condens. Matter* **3**, 4281–4287 (1991).
24. M. Ain, A. Menelle, B. M. Wanklyn, E. F. Bertaut, *J. Phys. Condens. Matter* **4**, 5327–5337 (1992).
25. H. Katsura, N. Nagaosa, A. V. Balatsky, *Phys. Rev. Lett.* **95**, 057205 (2005).
26. I. A. Sergienko, E. Dagotto, *Phys. Rev. B Condens. Matter Phys.* **73**, 094434 (2006).
27. M. Mostovoy, *Phys. Rev. Lett.* **96**, 067601 (2006).
28. S. Kibayashi, Y. Takahashi, S. Seki, Y. Tokura, *Nat. Commun.* **5**, 4583 (2014).
29. S. P. P. Jones *et al.*, *Nat. Commun.* **5**, 3787 (2014).
30. K. Cao, F. Giustino, P. G. Radaelli, *Phys. Rev. Lett.* **114**, 197201 (2015).
31. See the supplementary materials.
32. L. D. Landau, L. P. Pitaevskii, E. M. Lifshitz, *Electrodynamics of Continuous Media* (Butterworth-Heinemann, ed. 2, 1984).
33. L. Barron, *Molecular Light Scattering and Optical Activity* (Cambridge Univ. Press, ed. 2, 2004).
34. M. Saito, K. Taniguchi, T. Arima, *J. Phys. Soc. Jpn.* **77**, 013705 (2008).
35. T. A. Raybould *et al.*, *Phys. Rev. B* **94**, 035119 (2016).
36. V. Scagnoli *et al.*, *Science* **332**, 696–698 (2011).
37. T. Hoffmann, K. Kimura, T. Kimura, M. Fiebig, *J. Phys. Soc. Jpn.* **81**, 124714 (2012).
38. T. Kubacka *et al.*, *Science* **343**, 1333–1336 (2014).
39. R. Masuda, Y. Kaneko, Y. Tokura, Y. Takahashi, Replication data for: Electric field control of natural optical activity in a multiferroic helimagnet. Harvard Dataverse (2021); https://doi.org/10.7910/DVN/6X8OYQ.

NANOFLUIDICS

Atomic-scale ion transistor with ultrahigh diffusivity

Yahui Xue[1]†, Yang Xia[1]†, Sui Yang[1], Yousif Alsaid[1], King Yan Fong[1], Yuan Wang[1], Xiang Zhang[2]*

Biological ion channels rapidly and selectively gate ion transport through atomic-scale filters to maintain vital life functions. We report an atomic-scale ion transistor exhibiting ultrafast and highly selective ion transport controlled by electrical gating in graphene channels around 3 angstroms in height, made from a single flake of reduced graphene oxide. The ion diffusion coefficient reaches two orders of magnitude higher than the coefficient in bulk water. Atomic-scale ion transport shows a threshold behavior due to the critical energy barrier for hydrated ion insertion. Our in situ optical measurements suggest that ultrafast ion transport likely originates from highly dense packing of ions and their concerted movement inside the graphene channels.

B iological ion channels with atomic-scale selectivity filters not only allow extremely fast and precisely selective permeation of alkali metal ions but also behave as life's transistors, with the ability to gate their on-off responses to external stimuli so as to sustain important biological activities (1, 2). Efforts have been made to develop artificial pore structures to probe ion diffusion dynamics mimicking biological systems for fundamental understanding and applications in life science, filtration, and energy storage (3–5). Ion transport theory in a continuum framework has been developed with pores fabricated by conventional nanofabrication processes (6, 7), but these pore sizes exceed the diameters of hydrated ions, negating ion selectivity (8). Single-walled carbon nanotubes have emerged as interesting candidates for transporting ions (9). Nonetheless, these channels are mostly at the 1-nm scale, and they still exhibit poor selectivity given the typical angstrom-level size of hydrated ions and the measured diffusion coefficient, which is comparable to or even slower than that in bulk water (~1.9×10^{-9} m²/s) (10, 11). Recently, graphene layer structures have demonstrated their superiority over many others [e.g., graphene nanopores (12), metal-organic frameworks (13), and porous polymer (14)] in achieving elevated selectivity of monovalent metal ions with precisely controlled channel dimensions at the angstrom scale (15). However, these angstrom-scale channels preclude fast diffusion (on the order of 10^{-11} to 10^{-13} m²/s) because hydrated ions encounter steric resistance when entering a narrower channel space (16, 17). There is also a lack of gating ability at atomic scale to mimic the transistor-like behavior of biological channels. Although ion transport in so-called nanofluidic transistors and rectifiers has been reported in nanoscopic channels (18, 19), it still remains a grand challenge to realize gated ultrafast and simultaneously selective ion diffusion through atomic-scale channels to mimic the full functionalities of biological ones.

We report an atomic-scale ion transistor based on electrically gated graphene channels ~3 Å in height, possessing simultaneously ultrafast and selective ion transport that is two orders of magnitude faster than ion diffusion in bulk water. By applying electrical gating, the average surface potential on the graphene layer can be controlled, thus changing the energy barrier for ion intercalation into the channel. Our graphene channel device is made of a single flake of multilayer reduced graphene oxide (rGO) with interlayer spacing λ ≈ 0.45 nm (20). The device

[1]Nanoscale Science and Engineering Center, University of California, Berkeley, CA, USA. [2]Faculty of Science and Faculty of Engineering, University of Hong Kong, Hong Kong, China.
†These authors contributed equally to this work.
*Corresponding author. Email: president@hku.hk

Fig. 1. Experimental setup and ultrafast ion permeation. (**A**) Schematic of our atomic-scale graphene channel device with electric potential control using a three-electrode configuration. Gold is used as the gate electrode. Ag/AgCl in saturated KCl aqueous solution is used as the reference electrode (RE) and platinum wires as counter electrodes (CE). The arrows indicate the direction of ion flux driven by concentration gradient. Multiple graphene layers are electrically connected with the gate electrode in the out-of-plane direction (see fig. S6). (**B**) Schematics of "ON" and "OFF" states of ion permeation. Purple dots, alkali metal ions; light blue circles, hydration shells. In the "OFF" state, ions could be trapped from the previous round of ion intercalation as a result of cation-π interactions (17). (**C**) Experimental observation of potassium ion permeation through our graphene channels driven by a concentration gradient of 0.2 M at –1.2 V (ON state) and OCP (OFF state), respectively.
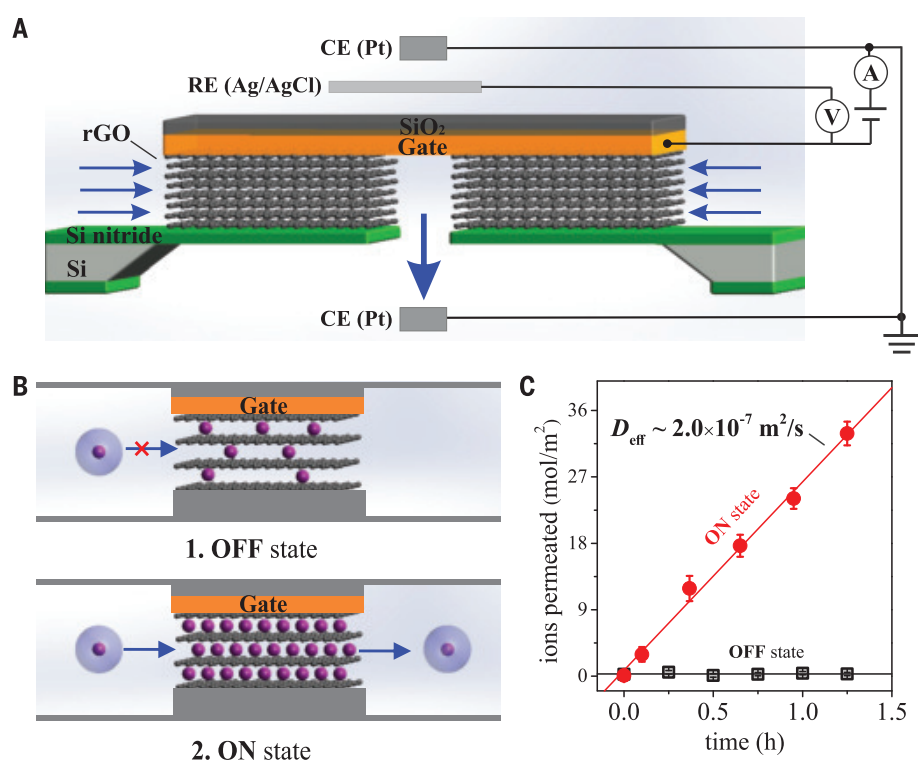
Fig. 2. Electrically controlled ion permeation.
(**A**) Potassium ion permeation through atomic-scale graphene channels of thickness t = 25 nm, driven by a concentration gradient of 0.2 M and at different electric potentials. Symbols are experimental data; lines are linear fittings.
(**B**) Permeation rate as a function of applied electric potential. Orange dots correspond to the experimental data in (A), obtained by the slopes of best linear fittings. Other symbols are data from two control experiments to exclude the leakage effect from the interfaces of the rGO flake with the silicon nitride substrate and the metal electrode, respectively (20).
(**C** and **D**) Potential control scheme (C) and ON-OFF responses (D) of ion permeation with switching potential between –1.2 V and OCP.
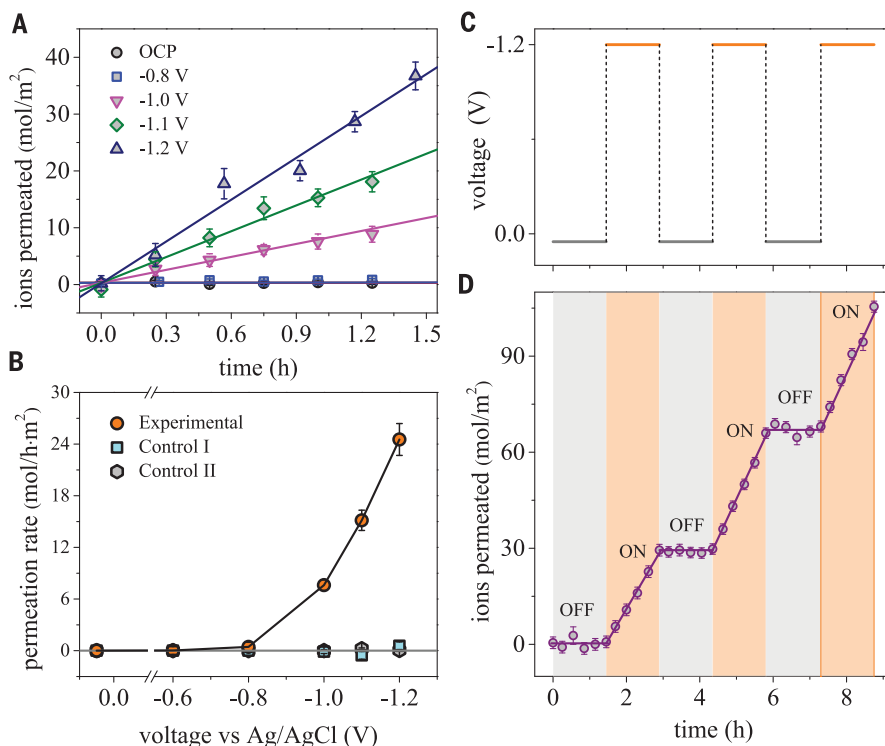
is supported on a free-standing silicon nitride membrane with a rectangular hole for ion flow and clamped with gate electrode and insulation layers for electric potential control (Fig. 1A). This configuration offers the advantages of intact layer structures (with diminished pinholes and defects relative to self-assembled graphene oxide laminates) for fundamental property investigation while preserving flexibility for scaling-up fabrication. Such two-dimensional graphene channels are clipped into rectangular shapes with length $L \approx 6$ μm and width $W \approx 20$ μm to facilitate ion permeation. The whole device as a membrane is clamped between two reservoirs (referred to as feed and permeate compartments), which are filled with highly conductive alkaline buffer electrolytes that can negate the electrophoretic effect and establish a 0.2 M concentration gradient of investigated cations [see (20) for detailed discussion on the electrophoretic effect].

To explore the possibility of controllable ion permeation through our graphene channels, we probed potassium ion diffusion purely driven by a concentration gradient (20). Because the channel size is smaller than the hydration diameters of alkali metal ions, it creates an intrinsic energy barrier that forbids ion entry in the open-circuit condition, indicating the OFF state for ion permeation (Fig. 1B, top). The electrical potential is applied to mimic the electric charge on the walls of biological channels. In this scenario, the hydration shell can be distorted or partially stripped off to

overcome the ion-entry energy barrier, enabling ion intercalation as capacitive charge (21). Beyond a percolation threshold as the intercalated ions accumulate inside the channels, permeable ion transport is initialized, indicating the ON state of ion permeation (Fig. 1B, bottom).

With this configuration, an ultrafast permeation of potassium ions is observed at –1.2 V (Fig. 1C, ON state), corresponding to a typical linear model driven by a concentration gradient and exhibiting an effective diffusion coefficient, $D_{eff} \approx 2.0 \times 10^{-7}$ m²/s, within graphene channels whose height is comparable to the selectivity filter diameter of biological channels. This diffusion coefficient is two orders of magnitude higher than that in bulk water and surpasses the intrinsic diffusion coefficient observed in biological channels (20). In contrast, potassium ions show negligible diffusion at the open-circuit potential (OCP) because of the steric hindrance of large-sized hydrated ions (Fig. 1C, OFF state).

To gain more insight into the effect of pore surface charge on ion permeation through atomic-scale channels, we investigated the permeation results under different electric potentials (Fig. 2A). Driven by a concentration gradient, the linear increase of ion concentration with time and with applied electric potential is observed beyond a threshold potential (–1.0 to –1.2 V). The higher the electric potential, the stronger the ion permeation. This suggests increased ion intercalation into atomic-scale graphene channels under strong

physical confinement with higher electrical driving force. Below the threshold (from OCP to –0.8 V), negligible change of ion permeation is observed. A plot of the potential-dependent permeation rate (Fig. 2B) shows that the threshold voltage behavior is evident, which indicates the onset electric potential to overcome the ion-entry energy barrier. Experimental evaluation of the energy barrier reveals a value of ~24 kJ/mol for potassium ions, which agrees reasonably well with predictions reported in the literature (15, 20, 22).

We further studied the switching capability of ionic permeation through the atomic-scale channels. We monitored the ion permeation by repeatedly switching the gate voltage below and beyond the threshold (Fig. 2C). It is seen that ion permeation happens linearly and rapidly with time at –1.2 V (beyond the threshold). When the potential is switched to OCP (below the threshold), the permeation immediately stops (Fig. 2D). This ON-OFF switchable sequence is repeatable over several cycles.

We explored the ion intercalation and packing density inside the channels under different electric potentials. Because the electronic structure of a material strongly affects its optical properties, we explored the intercalated charge density in our atomic-scale graphene channels by in situ optical reflection measurements (20). We observed enhanced optical reflection with respect to the open-circuit condition in the spectrum region from 450 to 900 nm beyond a threshold voltage (Fig. 3A

Fig. 3. Charge intercalation and ionic selectivity. (**A**) Relative optical reflection change induced by charge intercalation as a function of electric potential at wavelengths of 500, 600, 700, and 800 nm. R_0 is reflection intensity at OCP, and ΔR is relative change at different potential. Inset: schematic experimental setup for in situ optical reflection measurements. (**B**) Comparison of permeation rates of Li$^+$, K$^+$, and Cs$^+$ ions in atomic-scale graphene channels at –1.0, –1.1, and –1.2 V.
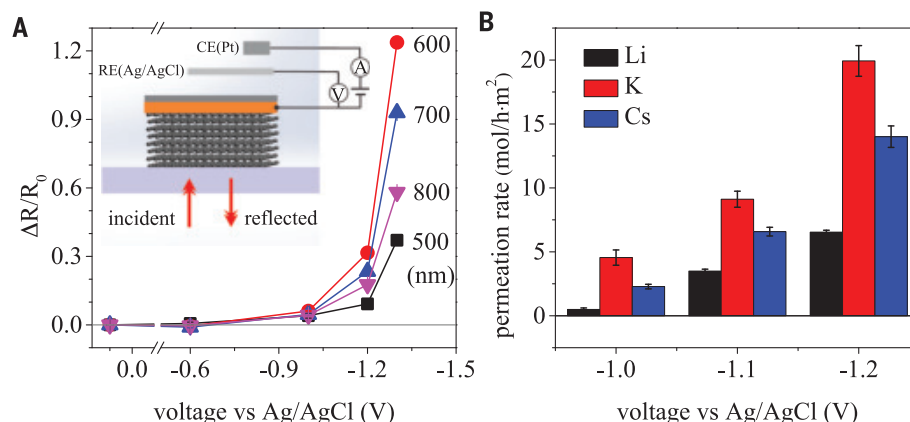
and fig. S14), consistent with our potential-dependent ion diffusion measurements. Such observation is attributed to the doping level increase as ions intercalate into the graphene channels. We modeled the optical reflection measurements by approximating the optical property of rGO with Fermi level–dependent multiple graphene layers (20). The modeling results capture well the overall trend of optical reflection change (fig. S16), revealing a charge density on the order of 1.8×10^{14} cm^{-2} at –1.2 V. Such a charge density is quite high and compares well with the measurements in the electrochemical charging of bilayer graphene (23) and ultrathin graphite (24).

We further probed how ion packing density promotes ultrafast ion diffusion. Note that ultimate confinement largely reduces the dielectric constant of aqueous solutions (25). The high packing density of charges observed in our graphene channels could impart strong Coulomb interaction among adjacent ions and promote their concerted movement with decreasing energy barrier (26, 27). We adopted a mean-field theoretical model to correlate the dependence of the effective ion diffusion coefficient on the charge density and their mutual interactions (20). For a charge density of 1.8×10^{14} cm^{-2}, the model yields an enhancement of the diffusion coefficient from 1.9×10^{-9} m^2/s to ~0.8×10^{-7} m^2/s, which reasonably agrees with our experimental results. The accelerated diffusion dynamics can be understood in terms of Coulomb interaction–induced concerted ion movement (27, 28). The theoretical underestimation may be caused by uncertainty due to modeling assumptions and the omission of factors such as the low friction of graphene surfaces (29) and phonon-enhanced ion diffusion (30).

Elevated selectivity of monovalent metal ions is also observed in our atomic-size graphene channels. To reasonably compare the permeation rates of different alkali metal ions, we filled the feed compartment with a mixture of LiCl, KCl, and CsCl, 0.2 M each in the same

solution. Fig. 3B shows the permeation rates of different cations in a potential range from –1.0 to –1.2 V (see also fig. S17). A selection ratio of K$^+$/Li$^+$ as high as 9.0 is achieved at –1.0 V. The overall permeation rates increase with the electric potential, consistent with the single–potassium ion measurements in Fig. 2A. It is also interesting to observe a constant selection sequence in the studied potential range—that is, K$^+$ > Cs$^+$ > Li$^+$—which resembles the selection sequence of biological potassium channels. This implies a controlling mechanism combined with ion dehydration and electrostatic interaction with ion-binding sites in our charged graphene channels (31), in contrast to previous artificial channels (13, 14).

Our atomic-scale ionic transistors can switch ion transport effectively and selectively. The ultrahigh ion diffusivity is explained by their dense packing and resultant concerted movement. The reversible and switchable ion permeation is attributed to the steric ion-entry energy barrier overcome by applied electrical potential. The synergetic contribution of atomic-scale confinement and surface potential renders the ion transport highly selective. This demonstration not only provides fundamental understanding of fast ion sieving in biological systems but also leads to highly controllable and ultrafast ion transport of relevance to ion batteries, seawater desalination, and medical dialysis.

REFERENCES AND NOTES

1. B. Eisenberg, *Fluct. Noise Lett.* **11**, 1240001 (2012).
2. F. J. Sigworth, *Nature* **423**, 21–22 (2003).
3. J. C. T. Eijkel, A. van den Berg, *Microfluid. Nanofluidics* **1**, 249–267 (2005).
4. J. R. Werber, C. O. Osuji, M. Elimelech, *Nat. Rev. Mater.* **1**, 16018 (2016).
5. C. C. Chen, L. Fu, J. Maier, *Nature* **536**, 159–164 (2016).
6. C. Duan, A. Majumdar, *Nat. Nanotechnol.* **5**, 848–852 (2010).
7. C. Dekker, *Nat. Nanotechnol.* **2**, 209–215 (2007).
8. H. B. Park, J. Kamcev, L. M. Robeson, M. Elimelech, B. D. Freeman, *Science* **356**, eaab0530 (2017).
9. J. Geng *et al.*, *Nature* **514**, 612–615 (2014).
10. F. Fornasiero *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17250–17255 (2008).
11. J. Wu, K. Gerstandt, H. Zhang, J. Liu, B. J. Hinds, *Nat. Nanotechnol.* **7**, 133–139 (2012).
12. R. C. Rollings, A. T. Kuan, J. A. Golovchenko, *Nat. Commun.* **7**, 11408 (2016).
13. H. Zhang *et al.*, *Sci. Adv.* **4**, eaaq0066 (2018).
14. P. Wang *et al.*, *Nat. Commun.* **9**, 569 (2018).
15. J. Abraham *et al.*, *Nat. Nanotechnol.* **12**, 546–550 (2017).
16. A. Esfandiar *et al.*, *Science* **358**, 511–513 (2017).
17. L. Chen *et al.*, *Nature* **550**, 380–383 (2017).
18. R. Karnik *et al.*, *Nano Lett.* **5**, 943–948 (2005).
19. L. Cantley *et al.*, *Nanoscale* **11**, 9856–9861 (2019).
20. See supplementary materials.
21. M. M. Hantel, T. Kaspar, R. Nesper, A. Wokaun, R. Kotz, *Electrochem. Commun.* **13**, 90–92 (2011).
22. Y. Yu *et al.*, *Nanoscale* **11**, 8449–8457 (2019).
23. M. Kühne *et al.*, *Nat. Nanotechnol.* **12**, 895–900 (2017).
24. W. Bao *et al.*, *Nat. Commun.* **5**, 4224 (2014).
25. L. Fumagalli *et al.*, *Science* **360**, 1339–1342 (2018).
26. D. A. Köpfer *et al.*, *Science* **346**, 352–355 (2014).
27. S. Kondrat, P. Wu, R. Qiao, A. A. Kornyshev, *Nat. Mater.* **13**, 387–393 (2014).
28. X. He, Y. Zhu, Y. Mo, *Nat. Commun.* **8**, 15893 (2017).
29. T. Mouterde *et al.*, *Nature* **567**, 87–90 (2019).
30. S. Muy *et al.*, *Energy Environ. Sci.* **11**, 850–859 (2018).
31. B. Hille, *Ionic Channels in Excitable Membranes* (Sinauer, ed. 2, 1992).

INDUCED SEISMICITY

A risk-based approach for managing hydraulic fracturing–induced seismicity

Ryan Schultz*, Gregory C. Beroza, William L. Ellsworth

Risks from induced earthquakes are a growing concern that needs effective management. For hydraulic fracturing of the Eagle Ford shale in southern Texas, we developed a risk-informed strategy for choosing red-light thresholds that require immediate well shut-in. We used a combination of datasets to simulate spatially heterogeneous nuisance and damage impacts. Simulated impacts are greater in the northeast of the play and smaller in the southwest. This heterogeneity is driven by concentrations of population density. Spatially varying red-light thresholds normalized on these impacts [moment magnitude (M_w) 2.0 to 5.0] are fairer and safer than a single threshold applied over a broad area. Sensitivity tests indicate that the forecast maximum magnitude is the most influential parameter. Our method provides a guideline for traffic light protocols and managing induced seismicity risks.

T
he injection of fluids into the subsurface has the potential to reactivate critically stressed faults (1). In particular, hydraulic fracturing has been recognized as a source of induced earthquakes (2), with potentially induced events as large as local magnitude (M_L) 5.7 causing substantial damage (3). Although these earthquakes are rare (4), the perceived risks of hydraulic fracturing have both caused public concern and impeded industry development (5, 6). Often, traffic light protocols have been used to manage the risks of induced earthquakes (table S1) (7, 8). Many unresolved questions remain about the efficacy of these protocols.

Recent work has better defined traffic light protocols (9–14), some within a risk-based framework (15). We define the red light as the threshold requiring immediate shut-in of the well that is causing the earthquakes. The red-light magnitude is thus chosen to minimize the risks of unacceptable shaking from post–shut-in seismicity (or continued operations). A magnitude threshold for the red light is simple to implement, and forecast modeling can tie those thresholds to risk-based targets of consequence (15). Hazards related to ground-motion nuisance and building damage are important considerations, particularly when hydraulic fracturing occurs in low-seismicity regions, where the population may be unfamiliar with or unprepared for earthquake shaking (2).

On the basis of this rationale, we developed a risk-based, red-light–threshold approach for the Eagle Ford shale play in Texas (16). The Eagle Ford formation has hosted some of the largest confirmed cases of hydraulic fracturing–induced earthquakes in the United States (17), albeit somewhat complicated by also having substantial extraction-related seismicity (18). Many of the requisite seismological datasets

Department of Geophysics, Stanford University, Stanford, CA, USA.
*Corresponding author. Email: rjs10@stanford.edu
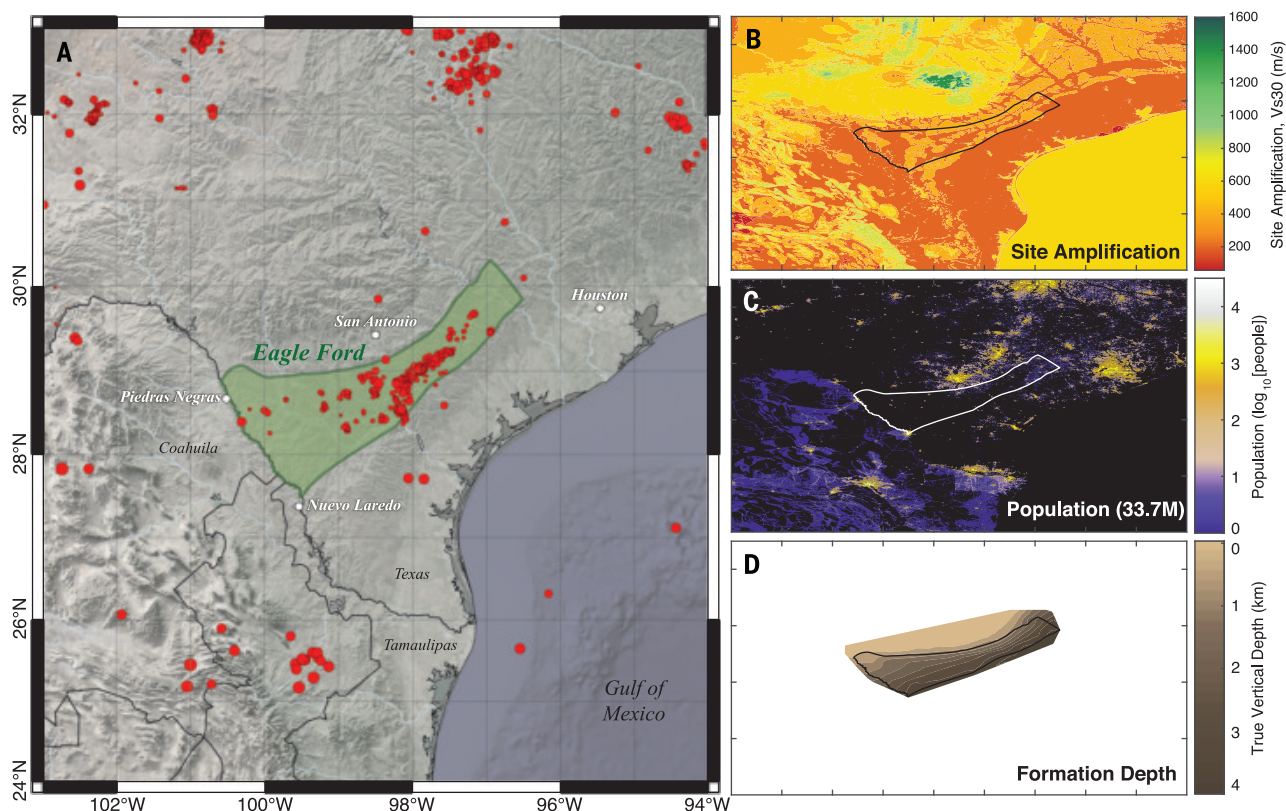
Fig. 1. Maps of spatial information, with the Eagle Ford boundaries. (**A**) Locations of earthquakes (red circles) and the boundaries of the Eagle Ford (green area) are shown alongside political boundaries and municipalities (white circles) for geographic context. (**B**) The same map bounds displaying a proxy for near–surface-site amplification (Vs30, scaled by color). (**C**) The same map bounds displaying the local population counts (log scaled by color, with black denoting zero population). (**D**) The same map bounds displaying the true vertical depth to the Eagle Ford formation (scaled by color).

are available (*19*), and the formation's slender outline transitions through a variety of population densities (Fig. 1 and figs. S1 to S3). For most of these reasons, the Eagle Ford provides an ideal example to illustrate probabilistic red-light thresholds to manage induced earthquake risks.

We gathered publicly available information on earthquake ground motion–prediction equations (*20*), proxies for site amplification of ground shaking (*21*), functions for translating ground-motion intensities to the probabilities of being felt or damaging (*22, 23*), and spatial distributions of population (*24*) to perform our analysis. Briefly, the workflow that we used simulates the largest induced earthquake trailing the shut-in of a hydraulic fracturing operation, including the event that initially triggered the red light (*25–27*). These trailing events are simulated over a range of red-light stopping magnitudes and all possible well locations within the Eagle Ford formation (Fig. 2A). From these simulations, we estimated the number of households that would likely be impacted by nuisance or damaging ground motions (Fig. 2B), similar to fatality estimates after an earthquake (*28*). Many of the input parameters to this problem have noteworthy variabilities (fig. S6), so we used a 300-trial Monte Carlo sample to represent earthquake

impacts as a statistical distribution (*15, 29*). Stopping at a larger-magnitude red light increases the number of impacted households; however, the number of households impacted can vary spatially (Fig. 2C), with the more densely populated areas in the northeast of the Eagle Ford (near Karnes City, San Antonio, and Houston) producing greater impacts. Although these impacts are spatially heterogeneous, the variabilities between individual realizations are far more pronounced (e.g., Fig. 2, B versus C). In this sense, we can design traffic light protocols that spatially vary to account for this controllable factor and then set risk tolerances that account for the uncontrollable factors.

This spatial heterogeneity of hydraulic fracturing–induced earthquake risks suggests that single-valued red-light magnitudes for an entire play are likely unfair: either too restrictive in rural areas (incurring unnecessary costs for operators) or too permissive near urban areas (potentially incurring inadequate safety). To address this issue, we preselected acceptable tolerances for nuisance or damage and then estimated the red-light stopping magnitude that would be required to meet them (e.g., Fig. 2C). For the purpose of demonstration, we chose an iso-nuisance tolerance of a 50% chance of 300,000 households

being impacted by a community decimal intensity (CDI) of 3. CDI is a questionnaire-based measure that quantifies the degree of earthquake shaking felt by a person (*30*), with levels 2 to 6 roughly corresponding to the subjective criteria of just felt, exciting, somewhat frightening, frightening, and extremely frightening, respectively. Our nuisance tolerance was chosen to keep the number of exciting (CDI 3) felt reports at a manageable level [because only ~0.2% of the population reports their experience (*23*)]. We chose an iso-damage tolerance of a 50% chance of 30 households being impacted at a damage state (DS) of 1. DS is a measure that quantifies the degree of building damage received from earthquake shaking, with levels 1 to 4 corresponding to damage that is slight or minor, moderate, extensive, and complete, respectively. This damage tolerance was chosen to minimize the chances that even cosmetic damage (DS 1) would occur. These tolerance choices are subjective value judgments. In practice, tolerances should be selected on the basis of input from all stakeholders (figs. S8 to S11).

We applied our methodology on the basis of these tolerances for nuisance and damage. Similar to the conceptual results, the number of households impacted by nuisance (given a
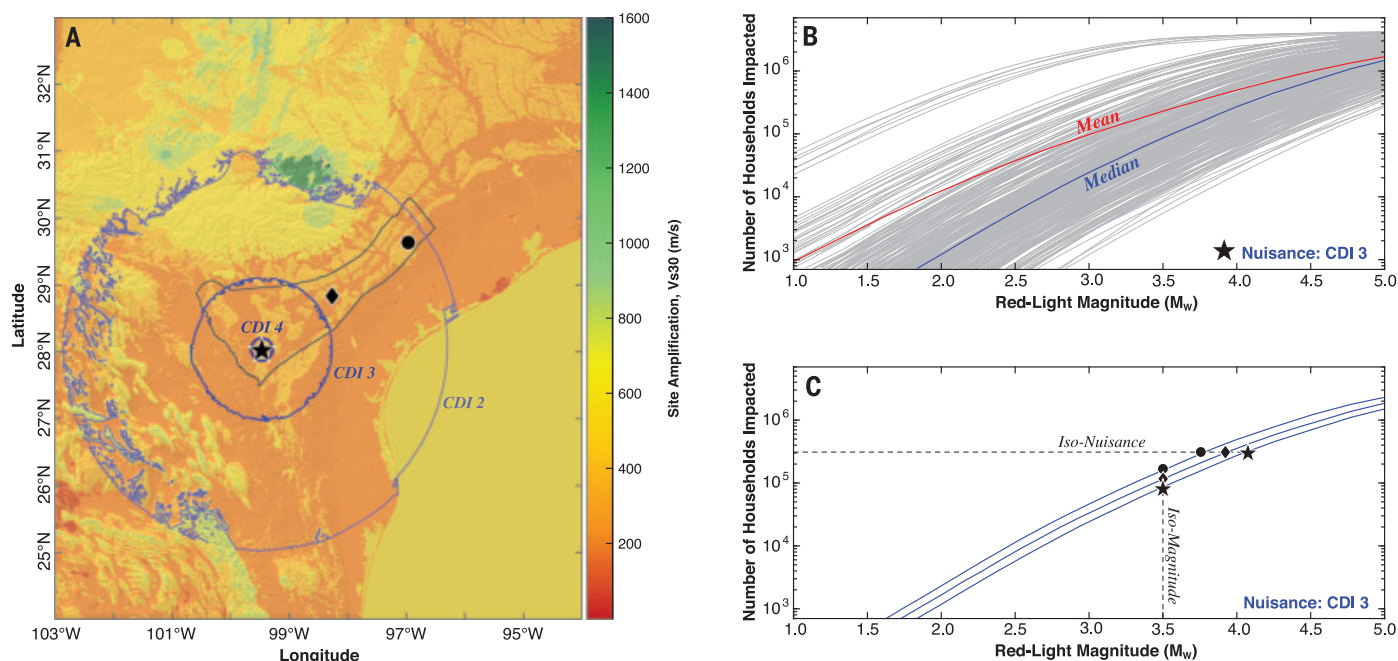
Fig. 2. The concepts of risk-based traffic light protocols. (**A**) Contours of 50% chance of felt ground motion at various nuisance degrees (blue circles and CDI labels) are estimated for a hypothetical induced earthquake (black star) in the Eagle Ford (black lines). A background map of site amplification (Vs30, colored area) provides context for deviations from circularity. The hypothetical ground motions can be computed for various geographic locations (star, diamond, and circle). Not visible, but the contours of 50% chance of encountering building DSs are also plotted underneath the star label. (**B**) Monte Carlo realizations of nuisance impacts at the CDI 3 level (gray lines) and the star location. The median (blue line) and mean curves (red line) may be produced from this distribution of realizations. (**C**) Median curves of nuisance impact (blue lines) for the three hypothetical induced earthquake locations on the left map (star, diamond, circle). Example iso-nuisance and iso-magnitude lines (dashed lines) depict the differing amount of median nuisance impact at the three locations.
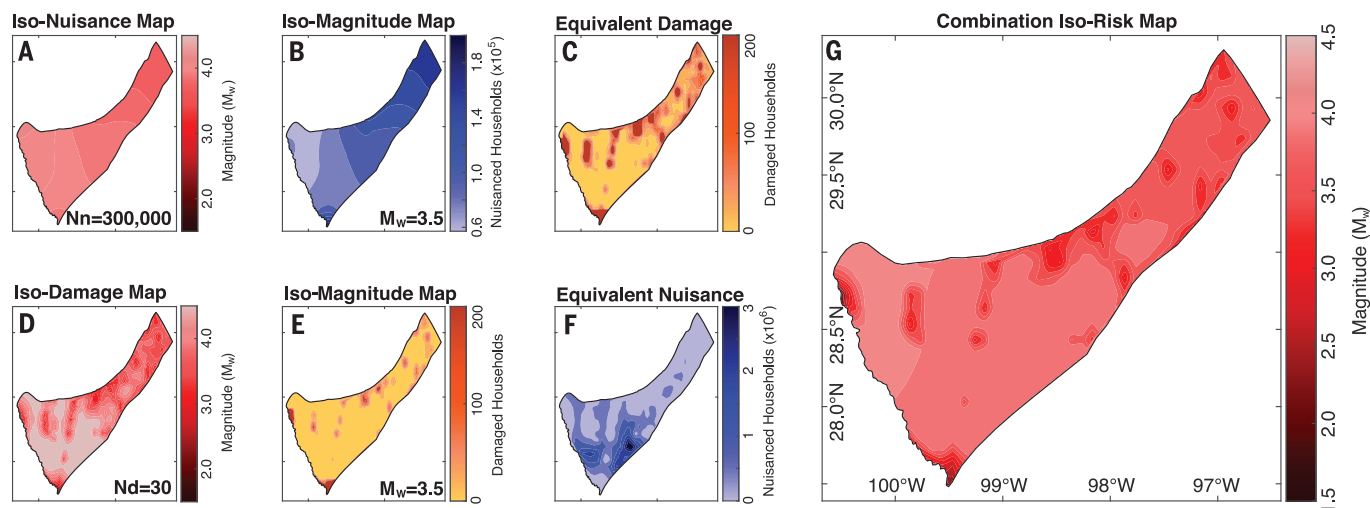
Fig. 3. Red-light impact maps. (**A**) Keeping a tolerance of a 50% chance of impacting 300,000 households at a nuisance level of CDI 3 constant, red-light stopping magnitudes are spatially estimated. (**B**) Keeping a red-light magnitude threshold of M_w 3.5, the median number of households impacted by nuisance at the CDI 3 level is estimated. (**C**) If the map produced in (A) is used as a red-light threshold, the median number of households impacted by damage at the DS 1 level is estimated (capped at 200). (**D**) Keeping a tolerance of a 50% chance of impacting 30 households at a damage level of DS 1 constant, red-light stopping magnitudes are spatially estimated. (**E**) Keeping a red-light magnitude threshold of M_w 3.5, the median number of households impacted by damage at the DS 1 level is estimated (capped at 100). (**F**) If the map produced in (D) is used as a red-light threshold, the median number of households impacted by nuisance at the CDI 3 level is estimated. (**G**) A combination of the two iso-risk maps (A and D), keeping the smaller magnitude at each grid pixel. Nd, number of households impacted by damage; Nn, number of households impacted by nuisance.

single-valued red-light threshold) increases toward the more densely populated northeast Eagle Ford formation (Fig. 3B). Nuisance impacts are spatially coherent, with a wavelength on the order of hundreds of kilometers, because moderate magnitude earthquakes are felt at these epicentral distances. Correspondingly, our iso-nuisance approach allows for higher red-light magnitudes of M_w 4.1 in the southwest that reduce to M_w 3.7 in the northeast (Fig. 3A). However, simply using the iso-nuisance magnitudes still produces heterogeneous damage impacts (Fig. 3C).

To explore these heterogeneous damage impacts, we repeated the methodology for damage. The number of households impacted by damage (given a single-valued red-light threshold) generally increases toward the northeast (Fig. 3E). However, this observation generally has a smaller wavelength, on the order of 1 to 10 km. This reflects the localized damage footprint of moderate magnitude earthquakes (31, 32). This creates "pockmark" features of lower red-light magnitudes (M_w ~3.0) in the iso-damage approach (Fig. 3D), which correspond to smaller-town locations (<300,000 households: e.g., Carrizo Springs, Cotulla, and Pearsall) that the iso-nuisance approach tends to miss. The southernmost and westernmost tips of the Eagle Ford have a lower red-light threshold of M_w ~2.0 where they encroach on the nearby Mexican cities of Nuevo Laredo and Piedras Negras. Selecting a red-light

magnitude on the basis of iso-damage alone produces heterogeneous nuisance impacts, especially in the sparsely populated southwestern area, where damage is unlikely (with red-light magnitudes up to M_w 5.0); however, allowing earthquakes of this magnitude has the potential to jeopardize the social license to operate (5).

To accommodate the shortcomings of either iso-risk approach, we produced a combination map. At every grid pixel, we take the smaller red-light magnitude between the iso-nuisance (Fig. 3A) and iso-damage maps (Fig. 3D). This combination map (Fig. 3G) ensures that neither the prescribed nuisance nor damage tolerances will be exceeded. In general, the sparsely populated rural areas tend to be controlled by the iso-nuisance tolerances, whereas iso-damage tolerances are more important in urban regions. The spatially averaged red-light magnitude is M_w 3.8 for this map.

We further tested the accuracy of our impact estimates and the level of restrictions our combination map would have imposed on operators. We retrospectively note that this combination map would have encountered four red-light earthquakes (e.g., M_w 4.8, 20 October 2011; fig. S12), on the basis of Eagle Ford's history of seismicity (17, 18). These four earthquakes were previously related to petroleum extraction (18). The statistics reported for these red-light earthquakes (~50% felt at CDI 3+) are very similar to the

tolerance we selected (fig. S13), providing some justification of our results. However, the retrospective analysis changes appreciably, depending on the tolerance for risk selected (figs. S8 to S11): Red lights for our most risk-tolerant and risk-averse choices range between 1 and 28 earthquakes (fig. S12). In regions where prior seismicity has occurred, these types of retrospective analyses could be used to inform risk tolerance choices.

Communities may have differing tolerances to nuisance or damage (15), and our approach provides useful guidelines for choosing traffic light thresholds. Our approach probabilistically defines red-light thresholds in a risk-based manner by combining nuisance and damage consequences while keeping magnitude-based red lights (which are simple to implement). Discretized versions of the combination map could also be used for further implementational simplicity (fig. S14). The yellow-light threshold (the point at which operator mitigation begins) could be defined based on the red-light threshold (e.g., M_w 2.0 lower than red) to minimize the number of green-to-red jumps encountered (15). Although this approach has been applied to the entire footprint of the Eagle Ford, only the susceptible regions that actually encounter earthquakes (33) would require management. Additionally, our approach could easily be adapted to hydraulic fracturing plays worldwide (2). Although this approach has been applied to hydraulic

Fig. 4. Sensitivity plots. (**A** and **B**) Sensitivity of perturbed input parameters on the output iso-risk results (A) for the iso-nuisance map and (B) for the iso-damage map. Values above the 95% confidence interval (dashed line) are notably influential parameters. Labels on the x axis are abbreviated as follows: b, *b*-value; dF1 and dF2, nuisance/fragility function covariances; dGM, ground-motion–prediction equation uncertainty; dM, forecast magnitude perturbation; dSA, site amplification perturbation; dPOP, population perturbation; dZ, earthquake depth perturbation.
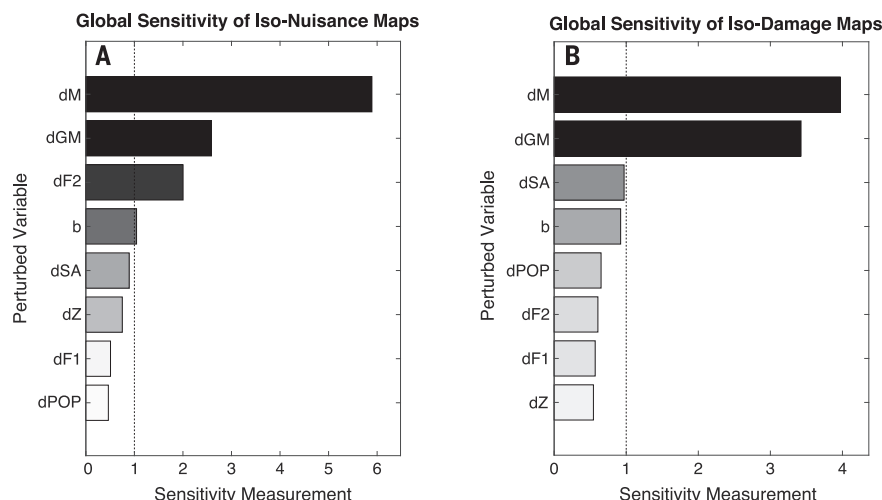
fracturing–induced seismicity, it would be applicable to induced earthquake risk in other settings, such as enhanced geothermal projects, or any other cases of induced seismicity with a quantifiable link between the causal operations and the induced sequence.

We acknowledge that this approach is subject to a number of variabilities and assumptions in our modeling (*15*). Many of these have the potential to change the simulated nuisance or damage impacts (e.g., Fig. 2B). Because of this, we used a statistical approach (*34*) to quantify the sensitivity of the output iso-nuisance and iso-damage results to the variable perturbations in the Monte Carlo analysis (Fig. 4). We found that the forecast maximum magnitude is the most influential parameter, followed by ground motion–prediction uncertainties, for both of the iso-risk maps. The sensitivity tests for the iso-nuisance map also suggest that perturbations to the nuisance likelihood function have an influence on the iso-nuisance map. All remaining parameters appear to have a minor influence on the resulting iso-maps. We interpret this ranking of parameter influences as a suggested order of research priorities; i.e., to better constrain red-light thresholds (and thus constrain induced earthquake risks), better models of the largest earthquakes after a well shut-in are needed the most (*27*). Accurate and regionally calibrated ground-motion predictions are also important, especially when considering damage impacts (Fig. 4B).

We have built a generalized approach to set red-light magnitude thresholds for managing the risks of hydraulic fracturing–induced seismicity. These thresholds are informed by simulated impacts of nuisance and damage. Furthermore, we applied this approach to the earthquakes induced in the Eagle Ford shale play. Although we have made particular risk tolerance choices in this paper to illustrate the methodology, communities, regulators, or operators would be expected to tailor tolerances according to their needs. Sensitivity tests indicate that better forecast models of earthquakes after the shut-in of a well and suitably calibrated ground motion–prediction equations are among the most important factors to improve. This approach to designing red-light thresholds in traffic light protocols will likely be useful for other cases of induced seismicity, such as enhanced geothermal systems.

REFERENCES AND NOTES

1. W. L. Ellsworth, *Science* **341**, 1225942 (2013).
2. R. Schultz et al., *Rev. Geophys.* **58**, e2019RG000695 (2020).
3. X. Lei, Z. Wang, J. Su, *Seismol. Res. Lett.* **90**, 1099–1110 (2019).
4. G. M. Atkinson et al., *Seismol. Res. Lett.* **87**, 631–647 (2016).
5. D. C. Smith, J. M. Richards, *Oil Gas Nat. Resour. Energy J.* **1**, 81–163 (2015).
6. H. Ghofrani, G. M. Atkinson, R. Schultz, K. Assatourians, *Seismol. Res. Lett.* **90**, 1420–1435 (2019).
7. J. J. Bommer et al., *Eng. Geol.* **83**, 287–306 (2006).
8. M. D. Zoback, *Earth Mag.* **57**, 38 (2012).
9. J. Douglas, H. Aochi, *Pure Appl. Geophys.* **171**, 1847–1858 (2014).
10. A. Mignan, D. Landtwing, P. Kästli, B. Mena, S. Wiemer, *Geothermics* **53**, 133–146 (2015).
11. A. Mignan, M. Broccardo, S. Wiemer, D. Giardini, *Sci. Rep.* **7**, 13607 (2017).
12. S. Baisch, C. Koch, A. Muntendam-Bos, *Seismol. Res. Lett.* **90**, 1145–1154 (2019).
13. C. Langenbruch, W. L. Ellsworth, J. U. Woo, D. J. Wald, *Geophys. Res. Lett.* **47**, e2019GL085878 (2020).
14. G. Cremen, M. J. Werner, *Nat. Hazards Earth Syst. Sci.* **20**, 2701–2719 (2020).
15. R. Schultz, G. Beroza, W. Ellsworth, J. Baker, *Bull. Seismol. Soc. Am.* **110**, 2411–2422 (2020).
16. US Energy Information Administration, "Updates to the EIA Eagle Ford Play Maps" (US Department of Energy, 2014); www.eia.gov/maps/pdf/eagleford122914.pdf.
17. S. L. Fasola et al., *Geophys. Res. Lett.* **46**, 12958–12967 (2019).
18. C. Frohlich, M. Brunt, *Earth Planet. Sci. Lett.* **379**, 56–63 (2013).
19. A. Savvaidis, B. Young, G. C. D. Huang, A. Lomax, *Seismol. Res. Lett.* **90**, 1702–1715 (2019).
20. G. Zalachoris, E. M. Rathje, *Earthq. Spectra* **35**, 1–20 (2019).
21. D. C. Heath, D. J. Wald, C. B. Worden, E. M. Thompson, G. M. Smoczyk, *Earthq. Spectra* **36**, 570–1584 (2020).
22. Federal Emergency Management Agency (FEMA). "Hazus-MH 2.1" in *Multi-Hazard Loss Estimation Methodology Technical and User Manuals* (Federal Emergency Management Agency, 2015).
23. R. Schultz, V. Quitoriano, D. J. Wald, G. Beroza, *Earthq. Spectra*, 10.1177/8755293020988025 (2021).
24. A. N. Rose, J. McKee, M. L. Urban, E. A. Bright, K. M. Sims, "LandScan 2018 High-Resolution Global Population Data Set" (No. LandScan 2018 High-Resolution Global Population Da; 005854MLTPL00; Oak Ridge National Laboratory, 2019).
25. N. J. van der Elst, M. T. Page, D. A. Weiser, T. H. Goebel, S. M. Hosseini, *J. Geophys. Res. Solid Earth* **121**, 4575–4590 (2016).
26. R. Schultz, G. Atkinson, D. W. Eaton, Y. J. Gu, H. Kao, *Science* **359**, 304–308 (2018).
27. J. P. Verdon, J. J. Bommer, *J. Seismol.* **25**, 301–326 (2021).
28. K. Jaiswal, D. Wald, *Earthq. Spectra* **26**, 1017–1037 (2010).
29. Materials and methods are available as supplementary materials.
30. D. J. Wald, V. Quitoriano, C. B. Worden, M. Hopper, J. W. Dewey, *Ann. Geophys.* **54**, 688–707 (2011).
31. R. E. Chase, A. B. Liel, N. Luco, B. W. Baird, *Earthquake Eng. Struct. Dynam.* **48**, 1365–1383 (2019).
32. B. W. Baird, A. B. Liel, R. E. Chase, *Earthq. Spectra* **36**, 1995–2018 (2020).
33. S. Pawley et al., *Geophys. Res. Lett.* **45**, 1786–1793 (2018).
34. J. Park, G. Yang, A. Satija, C. Scheidt, J. Caers, *Comput. Geosci.* **97**, 15–29 (2016).

QUANTUM COMPUTING

Parity-preserving and magnetic field–resilient superconductivity in InSb nanowires with Sn shells

M. Pendharkar[1]†, B. Zhang[2]†, H. Wu[2]†, A. Zarassi[2]†, P. Zhang[2]†, C. P. Dempsey[1], J. S. Lee[3], S. D. Harrington[4], G. Badawy[5], S. Gazibegovic[5], R. L. M. Op het Veld[5], M. Rossi[5], J. Jung[5], A.-H. Chen[6], M. A. Verheijen[5], M. Hocevar[6], E. P. A. M. Bakkers[5], C. J. Palmstrøm[1,3,4], S. M. Frolov[2]*

Improving materials used to make qubits is crucial to further progress in quantum information processing. Of particular interest are semiconductor-superconductor heterostructures that are expected to form the basis of topological quantum computing. We grew semiconductor indium antimonide nanowires that were coated with shells of tin of uniform thickness. No interdiffusion was observed at the interface between Sn and InSb. Tunnel junctions were prepared by in situ shadowing. Despite the lack of lattice matching between Sn and InSb, a 15-nanometer-thick shell of tin was found to induce a hard superconducting gap, with superconductivity persisting in magnetic field up to 4 teslas. A small island of Sn-InSb exhibits the two-electron charging effect. These findings suggest a less restrictive approach to fabricating superconducting and topological quantum circuits.

As we enter the era of intermediate-scale quantum circuits (1, 2), materials considerations come into renewed focus through their impact on quantum gate fidelity. The most successful solid-state approaches rely either on superconductors (3) or on semiconductors (4), with the future topological platform requiring a hybrid of the two (5). The search continues for the ultimate material capable of delivering low intrinsic decoherence. In this context, the effort to realize qubits based on Majorana zero modes that are expected to be topologically immune to decoherence (6–9) has accelerated the discovery of high-quality interfaces between superconducting metals and low-dimensional semiconductors (10–13).

Only a few superconductors have been explored for Majorana qubits, most notably aluminum, which is also the material of choice for transmon quantum processors (1). Among the advantages of aluminum are its self-limiting native oxide and a hard gap induced in proximate semiconductors (10, 12, 14, 15). Submicrometer islands of aluminum exhibit two-electron charging patterns: As a result of Cooper pairing, less total energy is required to add electrons in pairs than to add them one-by-one (16–23). This effect is a prerequisite for long quasiparticle stability times, crucial to both topological quantum computing and transmon qubits. Among the disadvantages of aluminum are a relatively small superconducting gap (with a temperature equivalent of 1 K) and a low critical magnetic field. This confines quantum computing based on superconducting aluminum to ultralow temperatures and places stringent constraints on the design of future topological qubits, which will require a precise balance of several energy scales (24).

Here, we present induced superconductivity in InSb nanowires (25, 26) with Sn shells. InSb is the highest electron mobility group III-V semiconductor with strong spin-orbit coupling (27) and large Landé g-factors in the conduction band (28). These are the primary ingredients of the Majorana recipe (29, 30), making InSb an optimal material for the

[1]Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA. [2]Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA. [3]California NanoSystems Institute, University of California, Santa Barbara, CA 93106, USA. [4]Materials Department, University of California, Santa Barbara, CA 93106, USA. [5]Eindhoven University of Technology, 5600 MB Eindhoven, Netherlands. [6]Univ. Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France.
*Corresponding author. Email: frolovsm@pitt.edu
†These authors contributed equally to this work.
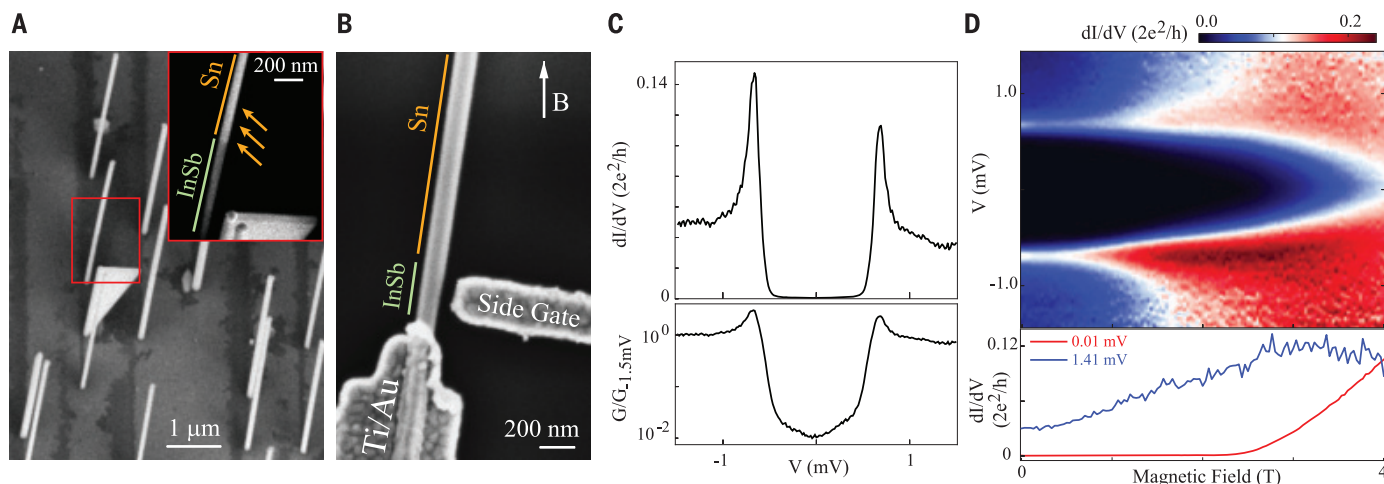
Fig. 1. Proximity-induced superconductivity with a hard gap. (A) Scanning electron micrograph (SEM) showing a triangular InSb flake that stood in the path of a beam of Sn atoms, shadowing the InSb nanowire standing behind. Dark streaks on the InSb substrate are also due to shadowing of the Sn beam by nanowires and flakes. The inset shows the direction of the Sn beam and indicates the shadowed and exposed segments of the nanowire. **(B)** SEM of an N-S device (device A) with a flake-shadowed Sn-InSb nanowire, Ti/Au contacts, and a side gate. The magnetic field is in the plane of the sample, as shown by the arrow. **(C)** Zero–magnetic field tunneling conductance spectrum of device A in linear scale (top) and logarithmic scale (bottom); $V_{BG} = 7.5$ V, $V_{SG} = -0.4$ V. **(D)** Top: Magnetic field evolution of the zero-field spectrum shown in (C). Bottom: Line cuts at two specific source-drain bias voltages. Measurements were performed at the base temperature of a dilution refrigerator, estimated to be 30 mK.

investigation of induced topological superconductivity (*31–35*).

We found that InSb nanowires coupled to tin exhibit a hard induced superconducting gap with a magnitude of up to 700 μeV. Superconductivity persists to magnetic fields up to 4 T for 15-nm-thick Sn shells. Islands of tin do exhibit 2*e*-periodic charging patterns, where *e* is the elementary charge. These re-

sults are obtained even though Sn and InSb have no epitaxial relationship, which was previously viewed as essential (*12, 14*).

Our first goal was to investigate electron tunneling into Sn through InSb in the normal metal-superconductor (N-S) configuration. For this we needed a nanowire with only one end covered by tin. We used the uncovered end to define a tunneling barrier and an N-contact.

To prevent unintentional damage to InSb that results from etching away part of the Sn shell, we used in situ nanoscale shadowing (*12, 23*). An InSb flake, standing in front of the nanowire, shadowed the bottom of the wire during deposition of Sn in ultrahigh vacuum (*36*) (Fig. 1A). We expect that junctions defined by wet etching or liftoff would yield similar results (*37, 38*).

Fig. 2. Superconducting tunneling. (**A**) SEM of device B (S-S device). Inset is a close-up of the shadow junction where the Sn shell is visible. (**B**) Differential conductance as a function of source-drain voltage bias *V* and back-gate voltage V_{BG}. The double arrows mark resonances 4Δ, 4Δ/2 (2Δ), and 4Δ/3. (**C**) Bottom: Differential resistance as a function of current bias *I* and V_{BG}. Top: Extracted switching current I_{SW} (black) and $I_{SW}R_N$ (red) as a function of V_{BG}. Measurements were performed at the base temperature of a dilution refrigerator, estimated to be 30 mK.
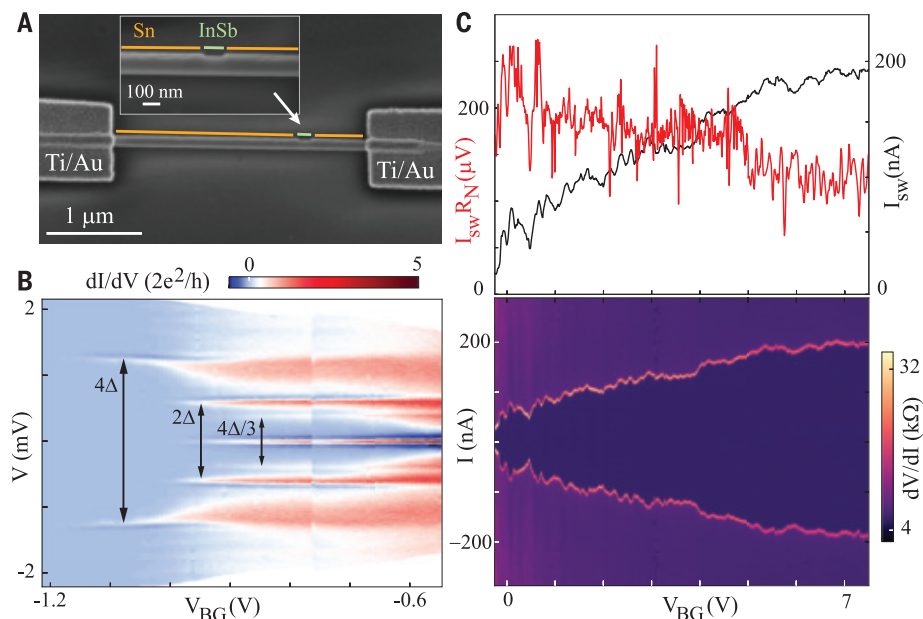
Fig. 3. Magnetic field dependence of the charging of tin islands. (**A**) SEM of device C showing the Sn island, two shadow junctions with bare InSb, side gates SG1 and SG2, and Ti/Au source-drain contacts, which cover the outside Sn segments and suppress superconductivity there. (**B** and **C**) 2*e* and 1*e* tunneling conductance resonances measured at *V* = −80 μV at B = 0 T and B = 1 T, respectively. (**D**) Magnetic field evolution of conductance along the dashed cut in (B). (**E** and **F**) *V* versus gate spectroscopy at B = 0 T and B = 1 T, respectively. The duration of data acquisition for each panel is on the order of 1 hour, and the parity pattern is reproducible over days. Measurements were performed at the base temperature of a dilution refrigerator, estimated to be 30 mK.
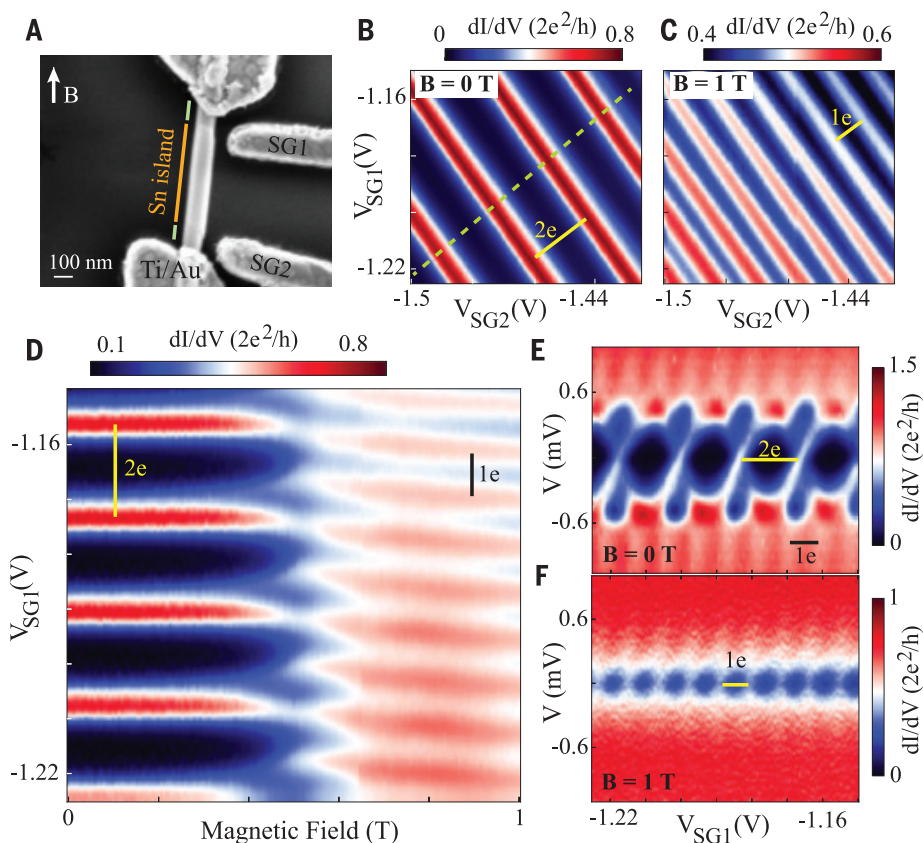
To prepare an N-S device, we positioned the tin-coated nanowire onto a doped Si/SiO$_x$ substrate, which was used as a back gate (BG) (Fig. 1B). A side gate (SG) was used to define and tune the tunneling barrier near the edge of the tin-free segment. The tunneling spectrum revealed a suppression in conductance by two orders of magnitude around zero bias (Fig. 1C). This so-called hard gap indicates the elimination of decoherence pathways caused by disorder and spurious subgap states. A superconducting tunneling peak is at ±680 μeV, which is comparable to the gap of tin. In a magnetic field, the hard gap was found to persist beyond 2 T, with the gap "softening" at higher fields and fully closing around 4 T (Fig. 1D). Magnetic field resilience is an indicator of a thin uniform shell; this resilience is another advantage of Sn, given that topological qubits, spin qubits, and some superconducting qubits operate at high magnetic fields. See fig. S1 for more data on the device presented in Fig. 1 (device A).

Next, we studied superconductor-superconductor (S-S) devices with both ends of the nanowire covered by tin and only a narrow break in the shell to define an InSb weak link (Fig. 2A). For this fabrication, we used a previously developed method of shadowing the Sn flux by criss-crossing nanowires (12). We first studied tunneling between two tin islands (Fig. 2B). We observed a smooth nanowire pinch-off void of accidental quantum dot states. Three finite-bias resonances were observed, marked 4Δ, 4Δ/2, and 4Δ/3 in Fig. 2B. This sequence is a manifestation of multiple Andreev reflection processes, which are characteristic of transparent S-S junctions. The positions of the resonances yield a value of Δ = 615 ± 10 μeV, which is somewhat smaller than the gap observed in the N-S device (Fig. 1C). At back-gate voltage $V_{BG} < -1$ V, only the 4Δ resonance was observed. We interpret this as the superconducting tunneling regime. Because the S-S tunneling resonance is a peak in current, it appears as a peak-dip structure in differential conductance.

The resonance at zero bias in Fig. 2B is the Josephson supercurrent. This effect is best studied in the current-bias configuration (Fig. 2C). The switching current (I_{sw}) from the superconducting to the normal state is a peak in differential resistance. I_{sw} decays smoothly with decreasing V_{BG}. The current-voltage characteristics are weakly hysteretic, as reflected in the asymmetry of I_{sw} in positive and negative current bias. In a magnetic field, the Josephson effect is observed up to 1.5 T and remains appreciable with sharp switching up to 0.5 T (fig. S4). This notable field range is a positive development for schemes that require coupling and decoupling of topologically superconducting islands at finite magnetic field for Majorana fusion or braiding (39, 40). Measurements on continuous-shell nanowires without shadow junctions yielded supercurrents in the range of 10 to 30 μA, corresponding to a critical current density of 2×10^6 A/cm^2 (fig. S7). The extracted products $I_{sw}R_N$ (where R_N is the normal-state resistance) are in the range of 125 to 225 μeV, which is of the same order of magnitude as the gap.

In Fig. 3 we present key findings on $2e$ charging of a tin island. The island was defined
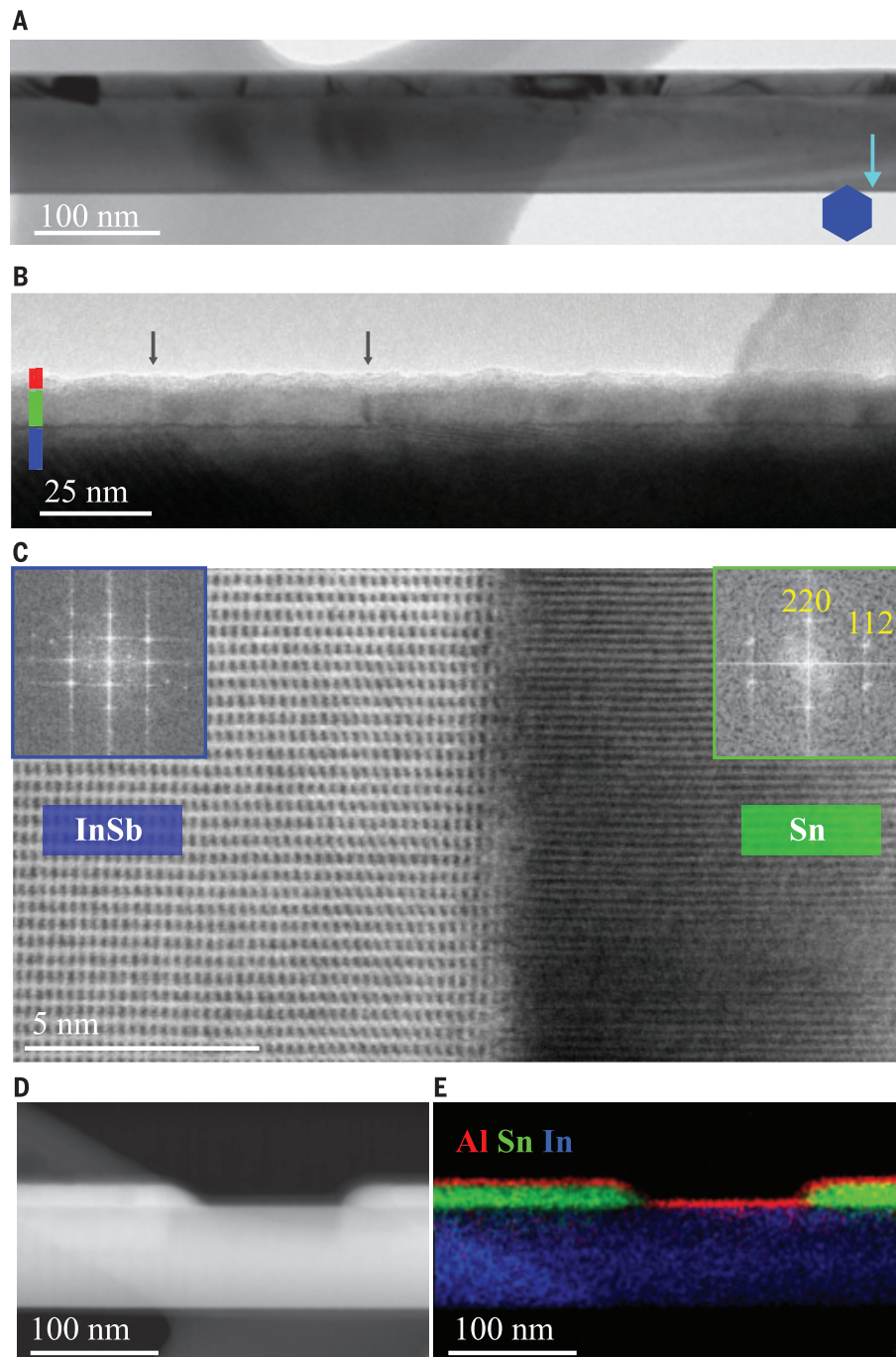
Fig. 4. Imaging of the InSb-Sn interface. (A) Side-view TEM image of an InSb-Sn nanowire along the ⟨112⟩ zone axis, showing a homogeneously thin shell. Arrow indicates imaging axis with respect to the hexagonal nanowire cross section. **(B)** Higher-magnification TEM of the AlO$_x$ (red)–Sn (green)–InSb (blue) stack. The Sn grain boundaries are highlighted by arrows. **(C)** High-resolution scanning TEM (HR-STEM) image of the Sn-InSb interface. The insets show Fourier transforms to the left and to the right of the interface. **(D)** High-angle annular dark-field STEM image of a shadow junction. **(E)** Energy-dispersive x-ray spectroscopy (EDX) elemental mapping of the shadow junction in (D). The Al-rich layer (red) corresponds to AlO$_x$; oxygen is not shown for clarity.

between two nanowire-shadow junctions in the N-S-N geometry (Fig. 3A). At zero magnetic field, we observed a single family of Coulomb peak resonances consistent with charging the entire island (Fig. 3B). At a finite magnetic field of 1 T, the frequency of Coulomb resonances doubled (Fig. 3, C and D). We attribute the data at zero field to 2e charging and the data at finite field to 1e charging. The transition from 2e to 1e is caused by the superconducting gap or the lowest subgap state dropping in energy below the charging energy, which we estimate to be 0.3 meV (Figs. 3, E and F). At finite magnetic field, it costs less energy to add electrons to the island one-by-one, whereas near zero field, owing to hard gap superconductivity, it is advantageous to add electrons in pairs. See figs. S3 to S6 for more data on the device shown in Fig. 3 (device C), as well as data from additional devices D, E, and F.

The two-electron charging effect is central to topological quantum computing because the states of a topological qubit are distinguished by even/odd island charge parity. If only 1e charging periodicity were observed, it would mean that despite a well-defined superconducting gap, electrons can be added to an island one at a time and the ability to distinguish the states of a topological qubit is lost. 1e periodicity is also detrimental for transmon qubits where single-electron tunneling is a decoherence mechanism.

Tin is an unusual material that has two different crystal phases with a phase transition at 13°C. The low-temperature α-Sn has a diamond cubic lattice, whereas the high-temperature β-Sn is tetragonal. The electronic properties of the two phases are very different. α-Sn is a semimetal that can also be a topological insulator in monolayer form (41–43), whereas β-Sn is a metal with a superconducting transition temperature of 3.7 K.

For tin on InSb nanowires, we assessed the structural properties and elemental distribution by transmission electron microscopy (TEM). TEM images reveal a polycrystalline Sn shell of uniform thickness around the InSb nanowire (Fig. 4A); the shell features grains of sizes 25 × 25 nm to 50 × 60 nm (Fig. 4B) (see also fig. S8). The Sn-InSb interface is abrupt and some Sn grains show an epitaxial relationship with InSb. The high-resolution annular bright-field scanning TEM image (Fig. 4C) shows a section of the interface where the {111} planes of zincblende InSb are aligned with lattice planes of a Sn grain with a lattice distance of 2.04 Å. This matches the {220} interplanar distance of β-Sn. This grain is one of 13 analyzed along the same nanowire (see fig. S9 and tables S1 and S2). Eleven of the grains were identified as β-Sn from the fast Fourier transform analysis of the interplanar distances (Fig. 4C, inset). Only two of those

β-Sn grains showed a preferential epitaxial relationship with InSb. In contrast, α-Sn is lattice-matched to InSb and can grow epitaxially (44). The predominantly β-Sn shell observed at room temperature by TEM is in agreement with superconductivity observed at low temperatures, which suggests that no phase transformation of Sn occurred upon device cooldown.

In addition to the uniform shell thickness, the nanowire shadow junctions used in S-S and N-S-N devices are sharp with Sn islands defined abruptly on each side of the junction (Fig. 4C). Energy-dispersive x-ray spectroscopy (EDX) confirms that the Sn islands are isolated from each other and no interdiffusion between Sn and In is detected (Fig. 4, D and E). A uniform 3-nm-thick AlO$_x$ passivation layer covers the entire nanowire; this layer maintains integrity and smoothness of the shell on a scale of at least 3 months. Oxidation at the Sn-InSb interfaces is not detected but cannot be fully excluded.

Our results illustrate that neither defect-free epitaxial wire-shell interfaces nor single vacuum-cycle growth of nanowire and shell are crucial requirements for the demonstration of hard-gap, field-resilient superconductivity and 2e charging. We conclude that the key components in attaining robust induced superconductivity are (i) removal of InSb native oxide using atomic hydrogen prior to Sn growth, followed by (ii) liquid nitrogen cooling of the nanowires during metal evaporation to produce a homogeneous ultrathin shell and (iii) immediate passivation of the wire-shell hybrid with a stable dielectric. Figure S10 illustrates that when the sample is allowed to warm up without AlO$_x$ passivation immediately after the cryogenic Sn shell growth, tin coagulates into discontinuous grains.

Without the need for epitaxial matching, many more superconductor-semiconductor combinations can be tried in searches for decoherence-free qubit materials (23, 38, 45–47). For Sn-InSb nanowires, we envision experiments on devices in the Majorana geometry in search of clear signatures of topological superconductivity, as well as characterization of coherence times of Sn-based transmon qubits.

REFERENCES AND NOTES

1. F. Arute et al., Nature 574, 505–510 (2019).
2. C. Figgatt et al., Nature 572, 368–372 (2019).
3. P. Krantz et al., Appl. Phys. Rev. 6, 021318 (2019).
4. C. Kloeffel, D. Loss, Annu. Rev. Condens. Matter Phys. 4, 51–81 (2013).
5. R. Lutchyn et al., Nat. Rev. Mater. 3, 52–68 (2018).
6. T. Hyart et al., Phys. Rev. B 88, 035121 (2013).
7. T. Karzig et al., Phys. Rev. B 95, 235305 (2017).
8. S. Plugge, A. Rasmussen, R. Egger, K. Flensberg, New J. Phys. 19, 012001 (2017).
9. J. P. Stenger, M. Hatridge, S. M. Frolov, D. Pekker, Phys. Rev. B 99, 035307 (2019).
10. P. Krogstrup et al., Nat. Mater. 14, 400–406 (2015).
11. J. Shabani et al., Phys. Rev. B 93, 155402 (2016).
12. S. Gazibegovic et al., Nature 548, 434–438 (2017).
13. A. Fornieri et al., Nature 569, 89–92 (2019).
14. W. Chang et al., Nat. Nanotechnol. 10, 232–236 (2015).
15. Ö. Gül et al., Nano Lett. 17, 2690–2696 (2017).
16. L. J. Geerligs, V. F. Anderegg, J. Romijn, J. E. Mooij, Phys. Rev. Lett. 65, 377–380 (1990).
17. M. T. Tuominen, J. M. Hergenrother, T. S. Tighe, M. Tinkham, Phys. Rev. Lett. 69, 1997–2000 (1992).
18. P. Lafarge, P. Joyez, D. Esteve, C. Urbina, M. H. Devoret, Phys. Rev. Lett. 70, 994–997 (1993).
19. T. M. Eiles, J. M. Martinis, M. H. Devoret, Phys. Rev. Lett. 70, 1862–1865 (1993).
20. P. Joyez, P. Lafarge, A. Filipe, D. Esteve, M. H. Devoret, Phys. Rev. Lett. 72, 2458–2461 (1994).
21. S. M. Albrecht et al., Nature 531, 206–209 (2016).
22. J. Shen et al., Nat. Commun. 9, 4801 (2018).
23. D. J. Carrad et al., Adv. Mater. 32, e1908411 (2020).
24. H. Pan, S. Das Sarma, Phys. Rev. Res. 2, 013377 (2020).
25. G. Badawy et al., Nano Lett. 19, 3575–3582 (2019).
26. S. R. Plissard et al., Nano Lett. 12, 1794–1798 (2012).
27. S. Nadj-Perge et al., Phys. Rev. Lett. 108, 166801 (2012).
28. H. A. Nilsson et al., Nano Lett. 9, 3151–3156 (2009).
29. R. M. Lutchyn, J. D. Sau, S. Das Sarma, Phys. Rev. Lett. 105, 077001 (2010).
30. Y. Oreg, G. Refael, F. von Oppen, Phys. Rev. Lett. 105, 177002 (2010).
31. V. Mourik et al., Science 336, 1003–1007 (2012).
32. M. T. Deng et al., Nano Lett. 12, 6414–6419 (2012).
33. J. Chen et al., Sci. Adv. 3, e1701476 (2017).
34. Ö. Gül et al., Nat. Nanotechnol. 13, 192–197 (2018).
35. M. T. Deng et al., Science 354, 1557–1562 (2016).
36. S. Gazibegovic et al., Adv. Mater. 31, e1808181 (2019).
37. S. T. Gill et al., Nano Lett. 18, 6121–6128 (2018).
38. S. A. Khan et al., ACS Nano 14, 14605–14615 (2020).
39. B. van Heck, A. Akhmerov, F. Hassler, M. Burrello, C. Beenakker, New J. Phys. 14, 035019 (2012).
40. D. Aasen et al., Phys. Rev. X 6, 031016 (2016).
41. Y. Xu et al., Phys. Rev. Lett. 111, 136804 (2013).
42. C.-Z. Xu et al., Phys. Rev. Lett. 118, 146402 (2017).
43. X. Zheng, J.-F. Zhang, B. Tong, R.-R. Du, 2D Mater. 7, 011001 (2019).
44. R. Farrow et al., J. Cryst. Growth 54, 507–518 (1981).
45. M. Bjergfelt et al., Nanotechnology 30, 294005 (2019).
46. N. A. Güsken et al., Nanoscale 9, 16735–16741 (2017).
47. T. Kanne et al., arXiv 2002.11641 [cond-mat.mes-hall] (26 February 2020).
48. M. Pendharkar et al., Data for "Parity-preserving and magnetic field–resilient superconductivity in InSb nanowires with Sn shells"; DOI: 10.5281/zenodo.4457486.

A widespread pathway for substitution of adenine by diaminopurine in phage genomes

Yan Zhou[1,2]†, Xuexia Xu[3,4,5,6]†, Yifeng Wei[7]†, Yu Cheng[3,4], Yu Guo[3,4], Ivan Khudyakov[8], Fuli Liu[9], Ping He[9], Zhangyue Song[10], Zhi Li[1], Yan Gao[1], Ee Lui Ang[7], Huimin Zhao[7,11]*, Yan Zhang[1,2]*, Suwen Zhao[3,4]*

DNA modifications vary in form and function but generally do not alter Watson-Crick base pairing. Diaminopurine (Z) is an exception because it completely replaces adenine and forms three hydrogen bonds with thymine in cyanophage S-2L genomic DNA. However, the biosynthesis, prevalence, and importance of Z genomes remain unexplored. Here, we report a multienzyme system that supports Z-genome synthesis. We identified dozens of globally widespread phages harboring such enzymes, and we further verified the Z genome in one of these phages, Acinetobacter phage SH-Ab 15497, by using liquid chromatography with ultraviolet and mass spectrometry. The Z genome endows phages with evolutionary advantages for evading the attack of host restriction enzymes, and the characterization of its biosynthetic pathway enables Z-DNA production on a large scale for a diverse range of applications.

Complete substitution of adenine (A) by diaminopurine (Z) in the genome of cyanophage S-2L was reported by Kirnos *et al.* in 1977 (*1*). Z forms three hydrogen bonds when base-paired with thymine (T) (*1, 2*). In contrast to other types of nucleobase modifications (*3*), Z is unusual in altering Watson-Crick base pairing, changing the physical, chemical, and mechanical properties of double-stranded DNA (*2, 4–6*). However, biochemical characterization of enzymes required for Z-genome synthesis has not been reported. Understanding the biosynthesis of the Z genome will shed light on the investigation of its prevalence. To date, cyanophage S-2L is the sole organism known to possess the Z base in its genome, although the Z base has been unambiguously identified in a carbonaceous meteorite and proposed as a nucleobase that could have been available for the origin of life (*7*).

The S-2L genome contains an open reading frame encoding PurZ, a homolog of adenylosuccinate synthetase (PurA) in the purine biosynthetic pathway (fig. S1) (*8*). PurA and adenylosuccinate lyase (PurB) catalyze the conversion of inosine 5′-monophosphate (IMP) to adenosine 5′-monophosphate (AMP) in many organisms (Fig. 1A) (*9*). It has been proposed that PurZ is involved in a similar reaction to supply Z nucleotides (Fig. 1A) (*3, 8*).

We tested this hypothesis using bioinformatic analysis followed by recombinant production and characterization of PurZs of various phages (Fig. 1B and table S1). Homology models were constructed for cyanophage S-2L PurZ (*Cp*PurZ) and its closest homolog, *Sinbacteraceae* bacterium phage contig *Sb*PurZ (Fig. 1C and table S1). The active sites of PurZs were then compared with that of *Escherichia coli* PurA (*Ec*PurA) (*10*). The substrates of PurA are IMP, guanosine 5′-triphosphate (GTP), and aspartate, and PurA catalyzes the transfer of the GTP γ-phosphate to IMP, followed by the displacement of the phosphate with aspartate to form adenylosuccinate (*9*). The catalytic residue Asp[13] in the PurA GDxxKG motif (where x is any amino acid) is replaced by a Ser residue in PurZ (Fig. 1, D and E, and fig. S2). Molecular modeling suggested that replacement of Asp with the less bulky Ser could accommodate an additional 2-amino group of the substrate (Fig. 1D) and could also alter the catalytic mechanism (fig. S3), because the Asp abstracts a proton from IMP in PurA (*9–11*). Multiple sequence alignment showed that although Asp[13] is highly conserved in PurA, the Asp→Ser substitution is a common feature shared by dozens of putative PurZ sequences (Fig. 1E and figs. S2 and S4). In addition, the conserved Arg[303] (*Ec*PurA numbering), which forms a hydrogen bond with the 2′-hydroxyl group of the IMP ribose, is replaced by an aliphatic Leu[278] in *Cp*PurZ or Leu[279] in *Sb*PurZ (Fig. 1E and fig. S2), consistent with the substrate being a deoxyribonucleotide instead of a ribonucleotide.

We screened a number of PurZ homologs from isolated phages or phage contigs in the metagenome and detected enzymatic activities for four versions [PurZs of *Acinetobacter* phage SH-Ab 15497 (*Ap*PurZ), *Salmonella* phage PMBT28 (*Sp*PurZ), and *Vibrio* phage phiVC8 (*Vp*PurZ) and *Sb*PurZ] (figs. S5 and S6 and tables S1 and S2), with 2′-deoxyguanosine 5′-monophosphate (dGMP), adenine 5′-triphosphate (ATP), and Asp as substrates (Fig. 1A), consistent with our homology model (Fig. 1, D and F).

Incubation of PurZs with saturating amounts of substrates resulted in time-dependent changes in the absorption spectra, indicating a reaction occurring at the nucleobase. Ultraviolet-visible (UV-Vis) difference spectra, obtained by subtracting the spectrum of the starting material from that of the product, exhibited a maximum (λ_{max}) at 287 nm (fig. S6, A to E), consistent with the conversion of dGMP (λ_{max}, 252 nm) (*12*) into a species resembling 2-amino-2′-deoxyadenosine 5′-monophosphate (dZMP) (λ_{max}, 247 and 280 nm) (*2*). During the reaction, a time-dependent increase in absorbance at 287 nm (ΔA_{287}) was accompanied by the proportionate release of phosphate, as detected by a colorimetric phosphomolybdate assay (Fig. 2A and fig. S6F) (*13*).

No reaction was observed with guanosine 5′-monophosphate (GMP) or IMP replacing dGMP, or with other ribonucleoside 5′-triphosphates (NTPs) replacing ATP, confirming its substrate specificity. The enzyme retained 38% activity when dIMP was used with ATP and Asp (Fig. 2A). The optimal pH of the reaction was ~9.5 (fig. S6G). Michaelis-Menten kinetics of PurZs was investigated using a continuous spectrophotometric assay monitoring ΔA_{287} (fig. S6H). For the four enzymes, the turnover rate (k_{cat}) ranges from 2.3 to 16.5 min^{-1}, and the Michaelis constant (K_m) ranges from 1.6 to 5.1 for dGMP and 4.1 to 21.1 μM for ATP (fig. S7 and table S3). The apparent K_m for dGMP is well below the previously reported intracellular concentration of dGMP in bacteria (~50 μM) (*14*). The k_{cat}/K_m for dIMP is much (15-fold) lower than that of the physiological substrate, dGMP.

The reaction intermediates and products of PurZs were confirmed by electrospray ionization tandem mass spectrometry (ESI-MS/MS) (Fig. 2, B to D, and fig. S6, I to L). Incubation of *Sb*PurZ with dGMP, ATP, and Asp resulted in the appearance of two new peaks at mass/charge ratios (*m/z*) of 426 and 461, corresponding to adenine 5′-diphosphate (ADP) and 2-aminodeoxyadenylosuccinate (ADAS), respectively (Fig. 2, B and C). Omission of Asp resulted in the appearance of one new peak at *m/z* = 426, which matched the mass of both ADP and the putative reaction intermediate 6-phosphoryl-dGMP (fig. S6I). No phosphate release was detected in the absence of Asp, consistent with its entrapment in 6-phosphoryl-dGMP (fig. S6J). The analogous intermediate, 6-phosphoryl-IMP, was observed in complex with the structures of

[1]Tianjin Key Laboratory for Modern Drug Delivery and High-Efficiency, Collaborative Innovation Center of Chemical Science and Engineering, School of Pharmaceutical Science and Technology, Tianjin University, Tianjin 300072, China. [2]Frontiers Science Center for Synthetic Biology (Ministry of Education), Tianjin University, Tianjin 300072, China. [3]iHuman Institute, ShanghaiTech University, Shanghai 201210, China. [4]School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China. [5]University of Chinese Academy of Sciences, Beijing 100049, China. [6]Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences, Shanghai 200061, China. [7]Singapore Institute of Food and Biotechnology Innovation, Agency for Science, Technology and Research (A*STAR), Singapore. [8]All-Russian Research Institute for Agricultural Microbiology, St. Petersburg 196608, Russia. [9]Department of Medical Microbiology and Immunology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China. [10]Biomedical Big Data Platform, SIAIS, ShanghaiTech University, Shanghai 201210, China. [11]Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.
†These authors contributed equally to this work.
*Corresponding author. Email: zhao5@illinois.edu (H.Z.); yan.zhang@tju.edu.cn (Y.Z.); zhaosw@shanghaitech.edu.cn (S.Z.)
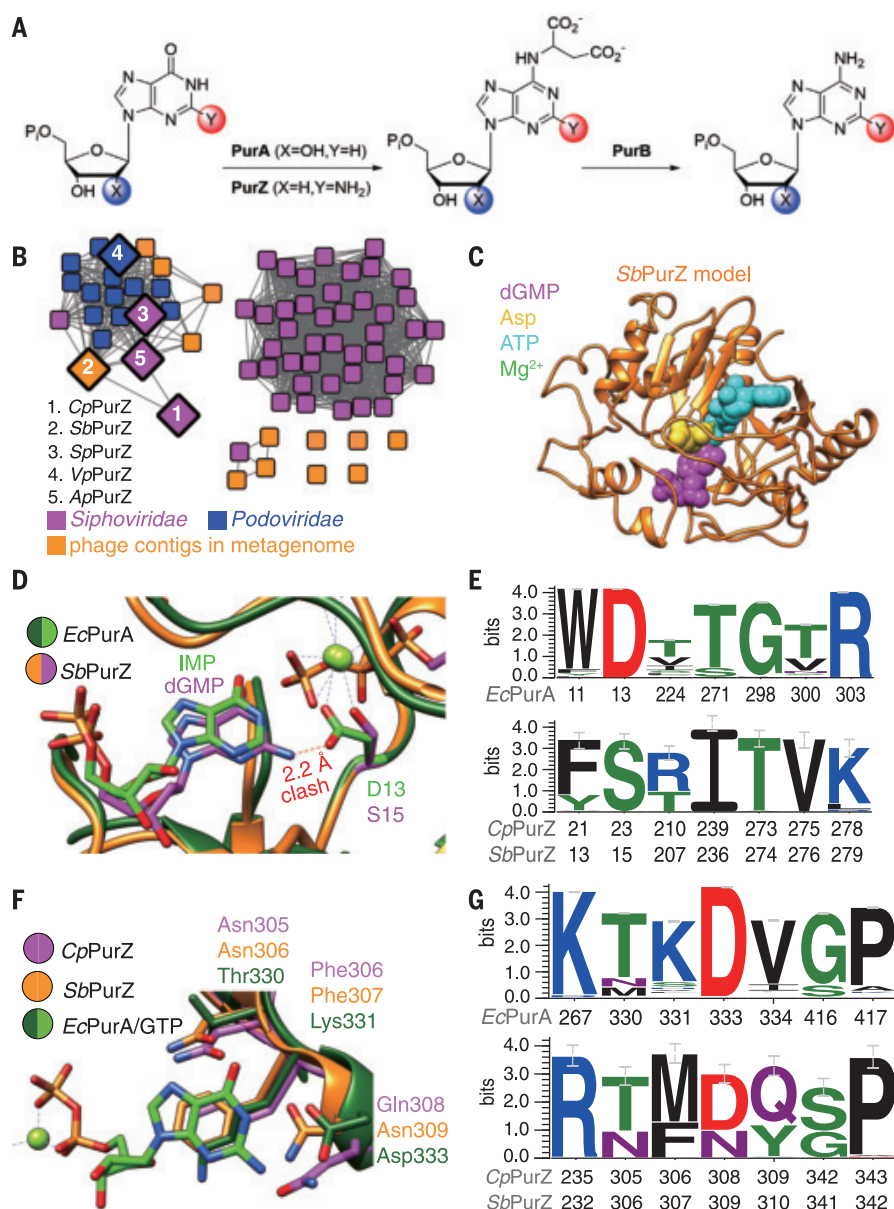
Fig. 1. Sequence and structural analyses of PurZs. (**A**) Analogous biosynthetic pathway for AMP and dZMP. (**B**) Sequence similarity network of putative PurZs. (**C**) Homology model of *Sb*PurZ. (**D**) The NMP binding site in *Sb*PurZ and *Ec*PurA. (**E**) Sequence logos of the NMP binding site. (**F**) Purine-interacting residues in the NTP binding sites of PurZs and *Ec*PurA. (**G**) Sequence logos of the NTP binding site. Residue numberings are from *Ec*PurA and *Cp*PurZ or *Sb*PurZ. D, aspartic acid; F, phenylalanine; G, glycine; I, isoleucine; K, lysine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine. Error bars indicate an approximate Bayesian 95% confidence interval.

*Ec*PurA crystallized in the presence of IMP, GTP, and hadacidin (*10*).

Because cyanophage S-2L does not encode homologs of PurB, we hypothesized that the subsequent conversion of ADAS to dZMP is catalyzed by PurB from the host (Fig. 3, A and B). Incubation of the *Sb*PurZ reaction products with recombinant *Ec*PurB (fig. S5E) resulted in the disappearance of the ADAS peak and appearance of a new peak at m/z = 345, corresponding to dZMP (Fig. 2D), demonstrating

that the bacterial PurB is competent for dZMP synthesis.

To investigate the substrate selectivity, mutagenesis of substrate-interacting residues in *Sb*PurZ to the corresponding residues in *Ec*PurA was performed. All mutations compromised the PurZ activity to varying degrees (table S4). Notably, the Ser[15]→Asp mutation abolished the activity of *Sb*PurZ, consistent with the role of Ser[15] in accommodating the 2-amino group of dGMP (Fig. 1D). The

Thr[274]→Gly mutant displayed a 271-fold increase in the K_m for Asp, consistent with its role in Asp binding (fig. S3). Mutants of the three residues interacting with adenine in ATP (Asn[306]→Thr, Phe[307]→Lys, and Asn[309]→Asp) resulted in an increase of the K_m for ATP from 89- to 450-fold (table S4), whereas none of them was active with GTP. Similarly, none of the mutants exhibited activity toward IMP. dZMP has to be phosphorylated to form 2-amino-2′-deoxyadenosine 5′-triphosphate (dZTP) before it can be incorporated into the phage genome by polymerase. This reaction was efficiently catalyzed by *Salmonella enterica* GMP kinase (*Se*GK), demonstrating that the function may be fulfilled by an enzyme from the bacterial host (Fig. 3B and figs. S5J and S8).

To investigate whether other enzymes are involved in Z-genome biosynthesis, we examined the genomic context of PurZ and identified a DNA polymerase, an HD domain–containing hydrolase-like enzyme, and a DUF550 domain–containing protein (Fig. 3A). HD domain enzymes appear in 20 phage genomes containing PurZ (Fig. 3A). These HD domain enzymes share a highly conserved metal- and ligand-binding pocket, despite their high sequence diversity (Fig. 3A and fig. S9). Recombinant *Cp*-, *Ap*-, and *Sp*HD enzymes (fig. S1; UniProt A0A2H5BHG9 and A0A2L0V156) were produced (fig. S5, F to H). Although they share only 28 to 33% sequence identity (Fig. 3A), all three enzymes exhibited 2′-deoxyadenine 5′-triphosphate triphosphohydrolase (dATPase) activity, catalyzing the hydrolysis of 2′-deoxyadenine 5′-triphosphate (dATP) to 2′-deoxyadenine (dA) and triphosphate, with the highest activity obtained using Co^{2+} as the divalent metal cofactor (Fig. 3C and figs. S10 to S12; apparent k_{cat} of 0.5 to 5.2 s^{-1} and apparent K_m of 6.5 to 74.8 μM). The apparent K_m of dATP is in the range of previously reported intracellular dATP concentrations in bacteria (*15*). The enzyme was highly specific for dATP, with much lower activities for hydrolysis of NTPs and other dNTPs (Fig. 3C). It also catalyzed the hydrolysis of 2′-deoxyadenine 5′-diphosphate (dADP) and 2′-deoxyadenine 5′-monophosphate (dAMP) into dA, releasing pyrophosphate and phosphate, respectively (figs. S11 and S12). Therefore, dATPase could facilitate Z-genome synthesis by specifically removing dATP and its precursor dADP from the nucleotide pool of the host (*16*), preventing the incorporation of A into the phage genome.

The DUF550 domain–containing protein is also of interest as it coexists with PurZ. Recombinant *Ap*DUF550 exhibited dATP and 2′-deoxyguanosine 5′-triphosphate (dGTP) pyrophosphohydrolase activity, catalyzing the hydrolysis of dATP/dGTP to pyrophosphate and dAMP/dGMP, respectively, with the highest activity obtained using Co^{2+} as the divalent metal cofactor and little or no activities for

Fig. 2. Substrate specificities and mass spectrometry assay results of *Sb*PurZ.
(**A**) Substrate specificity of *Sb*PurZ. Data are means ± SEM (*N* = 3 replicates). (**B**) Complete reaction with *Sb*PurZ, dGMP, ATP, and Asp. (**C**) Reaction with Asp, dGMP, and ATP, omitting *Sb*PurZ. (**D**) The two-step reaction coupling *Sb*PurZ and *Ec*PurB.
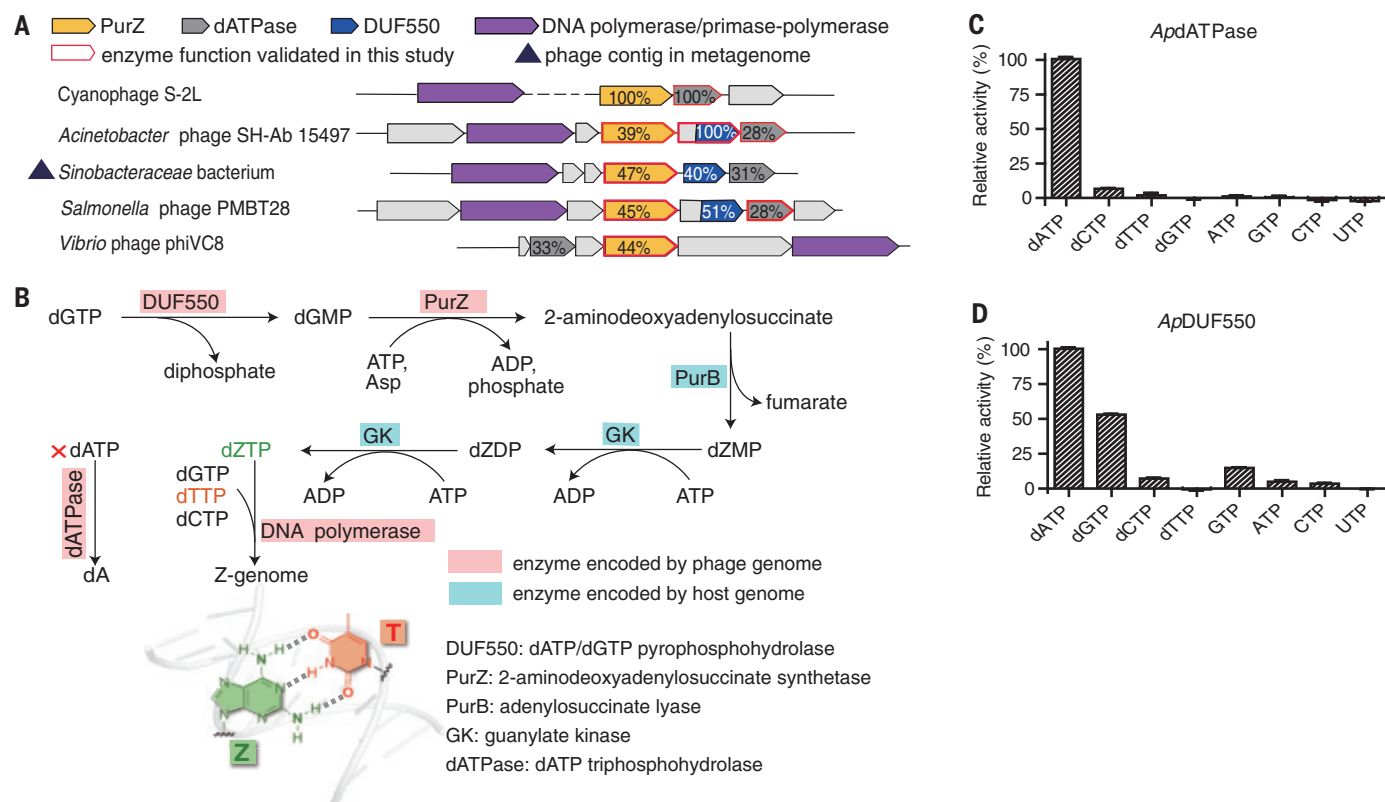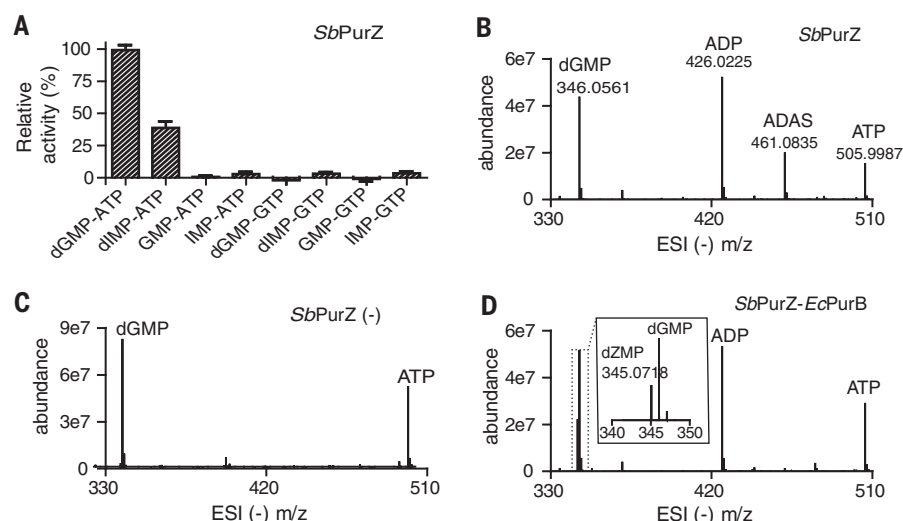
Fig. 3. Proposed biosynthetic pathway of phage Z genome. (**A**) Genome neighborhood of PurZ in phages. (**B**) Biosynthetic pathway of Z genome. (**C** and **D**) Substrate specificities of *Ap*dATPase (C) and *Ap*DUF550 (D). Data are means ± SEM (*N* = 3 replicates).

NTPs and pyrimidine dNTPs (Fig. 3D and fig. S13). Therefore, this DUF550-containing enzyme could function to supply dGMP as the substrate for PurZ, elevating dZTP level but depleting dATP to further promote Z incorporation (Fig. 3B).

Identification of PurZ and the other genes involved in dZTP biosynthesis and incorporation allowed us to investigate the occurrence of the Z genome in nature. Predicted PurZ sequences include 60 from bacteriophage isolates, with 13 from phage contigs in metagenomes (Fig. 1B and data S1). PurZ-containing bacteriophages largely belong to the families *Podoviridae* and *Siphoviridae* (Fig. 1B). The possibility that these phages contain modified purines in DNA might have been overlooked in the postgenomic era, when chemical determination of nucleotide content or base composition was no longer routine. We developed a protocol to confirm the Z incorporation and chose one of these phages, lytic *Acinetobacter* phage SH-Ab 15497, for further investigation (*17*). The phage DNA was prepared and digested with a combination of DNase I, phosphodiesterase I, and alkaline phosphatase, followed by liquid chromatography (LC)–UV spectrometric and LC-MS/MS analyses (Fig. 4, A and B). Compared with "normal" DNA (A-DNA), the digest products of the phage DNA and Z-DNA [the positive
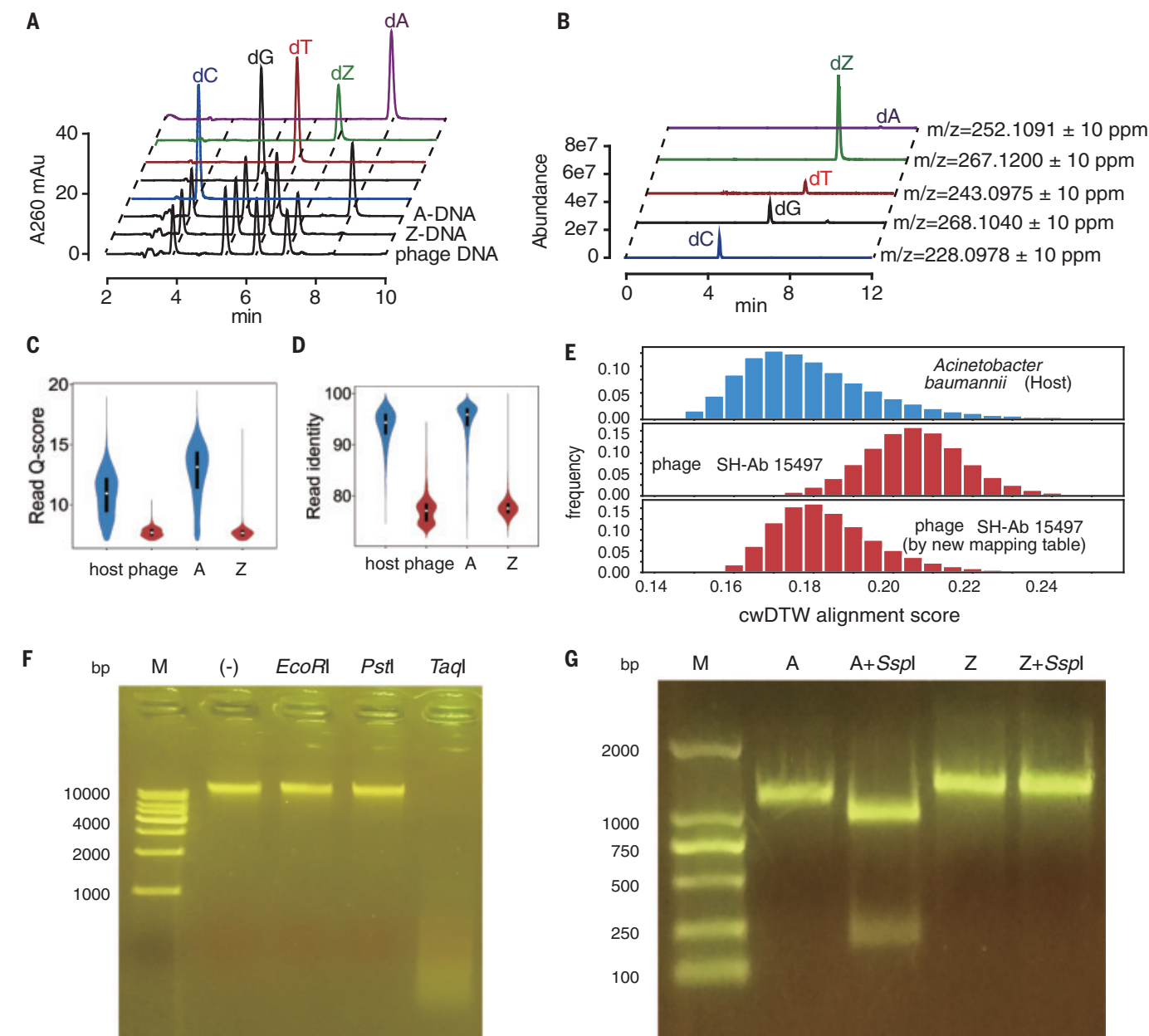
Fig. 4. Validation of Z incorporation in the genome of *Acinetobacter* phage SH-Ab 15497. (**A** and **B**) LC-UV detection (A) and extracted-ion chromatogram (B) of dZ from the phage DNA hydrolysis products. mAu, milli-absorbance unit; ppm, parts per million. (**C** and **D**) Read Q-score (C) and read identities (D) of the host, phage, and A- and Z-containing *Ap*PurZ PCR products with corresponding next-generation sequencing results. (**E**) cwDTW score distribution of host and phage reads. (**F**) Restriction enzyme digestions of the phage genomic DNA. Lane M, markers (in base pairs); -, control lane. (**G**) SspI can digest A-containing but not Z-containing *Ap*PurZ PCR products.

control; i.e., the polymerase chain reaction (PCR) product of *Ap*PurZ using Q5 polymerase, with dZTP replacing dATP in the dNTP mixture (fig. S14)] have negligible amounts of dA (Fig. 4A). Using extinction coefficients to quantify the UV peaks yields the molar ratios of deoxynucleosides that fit Chargraff's rule, with the dZ-to-dT ratio at 0.99 and the dC-to-dG ratio at 1.05, respectively. Positive ion masses matching the dZ nucleoside and Z base were detected (Fig. 4B and fig. S15). These results suggest that Z nearly com-

pletely, if not completely, replaces A to pair with T in the genome of SH-Ab 15497, that the Z-containing genome is not restricted to cyanophage S-2L, and that its occurrence may be more widespread than previously appreciated.

Because our phage DNA extractions always contain varied amounts of host DNA fragments, we performed nanopore sequencing of the crude extract to show that the quality [quantified by Q-score, read identity, and continuous wavelet dynamic time warping

(cwDTW) alignment score] of phage reads resembles that of the Z-DNA control, whereas the quality of host reads resembles that of "normal" DNA (Fig. 4, C to E, and figs. S16 and S17) (*18–20*). The nanopore signals, assigned to ~100 million reads of phage and host DNA each, allowed us to statistically analyze our data and comfortably conclude that the phage DNA is nearly free of adenine, whereas Z is likely not incorporated into the host DNA.

The function of many bacteriophage DNA modifications is to evade the restriction enzyme

attacks of hosts (3). The genomic DNA of S-2L has been reported to be resistant to most restriction enzymes (21, 22). To confirm that this is generally true for Z genomes, we investigated the susceptibility of the genomic DNA of SH-Ab 15497 and the Z-PCR products to digestion by restriction enzymes. We observed that Z-substituted DNAs from both sources were resistant to digestion by all the enzymes with recognition sites containing one or more As, including two close homologs of the host endogenous DNA restriction enzymes (EcoRI and PstI). The only exception is TaqI, which is known to be able to tolerate various DNA modifications (21). By contrast, HaeIII and Sau96I both recognize GC-only sequences and can readily digest the phage DNA and the Z-DNA control (Fig. 4, F and G; fig. S18; table S5). Therefore, one evolutionary advantage conferred by the Z genome may be the ability to escape restriction digestion in a variety of bacteria. No distinct degradation was visible for phage DNA digested with Sau3AI, for which the phage DNA contains more than 200 recognition sites, consistent with the apparently complete replacement of A by Z in the phage genome (fig. S18). We further show that PurZ-containing phages are widely distributed on Earth (fig. S19).

Here we have reported the Z-genome biosynthetic system, which could facilitate production of Z-substituted DNA for a wide range of emerging applications, such as DNA origami (23) and DNA-based data archiving, which has enormous potential because of its high storage capacity (24). Incorporation of the Z-genome biosynthetic enzymes into engineered phages may extend their host range and potency for applications in phagotherapy [which has been used successfully in clinical settings and has saved lives from infection with multidrug-resistant *A. baumannii* (25) or *Mycobacterium abscessus* (26)], for food preservation (27), and for environmental purposes. Scientists have long dreamed of increasing the diversity of bases (28–30); our work shows that nature has already come up with a way to do that. Considering that the Z base was discovered in a meteorite, our work may spark interest in interdisciplinary research on the origins of life and astrobiology (31).

REFERENCES AND NOTES

1. M. D. Kirnos, I. Y. Khudyakov, N. I. Alexandrushkina, B. F. Vanyushin, *Nature* **270**, 369–370 (1977).
2. I. Y. Khudyakov, M. D. Kirnos, N. I. Alexandrushkina, B. F. Vanyushin, *Virology* **88**, 8–18 (1978).
3. P. Weigele, E. A. Raleigh, *Chem. Rev.* **116**, 12655–12687 (2016).
4. C. Cheong, I. Tinoco Jr., A. Chollet, *Nucleic Acids Res.* **16**, 5115–5122 (1988).
5. J. Sági, E. Szakonyi, M. Vorlíčková, J. Kypr, *J. Biomol. Struct. Dyn.* **13**, 1035–1041 (1996).
6. M. Cristofalo et al., *Biophys. J.* **116**, 760–771 (2019).
7. M. P. Callahan et al., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13995–13998 (2011).
8. P. Marliere et al., Patent WO2003093461 (2003).
9. R. B. Honzatko, H. J. Fromm, *Arch. Biochem. Biophys.* **370**, 1–8 (1999).
10. J. Y. Choe, B. W. Poland, *Biochemistry* **38**, 6953–6961 (1999).
11. C. Kang, N. Sun, R. B. Honzatko, H. J. Fromm, *J. Biol. Chem.* **269**, 24046–24049 (1994).
12. M. J. Cavaluzzi, P. N. Borer, *Nucleic Acids Res.* **32**, e13 (2004).
13. G. Hua et al., *Anal. Biochem.* **547**, 26–36 (2018).
14. B. D. Bennett et al., *Nat. Chem. Biol.* **5**, 593–599 (2009).
15. B. R. Bochner, B. N. Ames, *J. Biol. Chem.* **257**, 9759–9769 (1982).
16. B. L. Greene et al., *Annu. Rev. Biochem.* **89**, 45–75 (2020).
17. Y. Hua et al., *Acta Biochim. Biophys. Sin.* **51**, 1079–1081 (2019).
18. J. J. Kasianowicz, E. Brandin, D. Branton, D. W. Deamer, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13770–13773 (1996).
19. L. Xu, M. Seki, *J. Hum. Genet.* **65**, 25–33 (2020).
20. R. Han, Y. Li, X. Gao, S. Wang, *Bioinformatics* **34**, i722–i731 (2018).
21. M. Szekeres, A. V. Matveyev, *FEBS Lett.* **222**, 89–94 (1987).
22. A. Chollet, E. Kawashima, *Nucleic Acids Res.* **16**, 305–317 (1988).
23. P. W. Rothemund, *Nature* **440**, 297–302 (2006).
24. L. Ceze, J. Nivala, K. Strauss, *Nat. Rev. Genet.* **20**, 456–466 (2019).
25. R. T. Schooley et al., *Antimicrob. Agents Chemother.* **61**, e00954-17 (2017).
26. R. M. Dedrick et al., *Nat. Med.* **25**, 730–733 (2019).
27. Z. D. Moye, J. Woolston, A. Sulakvelidze, *Viruses* **10**, 205 (2018).
28. Y. Zhang et al., *Nature* **551**, 644–647 (2017).
29. S. Hoshika et al., *Science* **363**, 884–887 (2019).
30. D. A. Malyshev et al., *Nature* **509**, 385–388 (2014).
31. J. Gollihar, M. Levy, A. D. Ellington, *Science* **343**, 259–260 (2014).

NONCANONICAL GENOME

A third purine biosynthetic pathway encoded by aminoadenine-based viral DNA genomes

Dona Sleiman[1]†, Pierre Simon Garcia[2,3]†, Marion Lagune[1], Jerome Loc'h[4], Ahmed Haouz[5], Najwa Taib[2,6], Pascal Röthlisberger[7], Simonetta Gribaldo[2], Philippe Marlière[8], Pierre Alexandre Kaminski[1]*

Cells have two purine pathways that synthesize adenine and guanine ribonucleotides from phosphoribose via inosylate. A chemical hybrid between adenine and guanine, 2-aminoadenine (Z), replaces adenine in the DNA of the cyanobacterial virus S-2L. We show that S-2L and *Vibrio* phage PhiVC8 encode a third purine pathway catalyzed by PurZ, a distant paralog of succinoadenylate synthase (PurA), the enzyme condensing aspartate and inosylate in the adenine pathway. PurZ condenses aspartate with deoxyguanylate into dSMP (N6-succino-2-amino-2′-deoxyadenylate), which undergoes defumarylation and phosphorylation to give dZTP (2-amino-2′-deoxyadenosine-5′-triphosphate), a substrate for the phage DNA polymerase. Crystallography and phylogenetics analyses indicate a close relationship between phage PurZ and archaeal PurA enzymes. Our work elucidates the biocatalytic innovation that remodeled a DNA building block beyond canonical molecular biology.

Bacteriophage genomes contain many modified nucleotides that are enzymatically synthesized and then incorporated by polymerization. The most conspicuous is 2-aminoadenine (hereafter referred to as Z), which was found in *Synechococcus* phage S-2L (1) and has also been detected in meteorites, suggesting a prebiotic existence (2). Z completely replaces the canonical adenine in S-2L DNA (3), increasing its thermostability because of a third hydrogen bond in the pair Z:T (4) and altering the conformational prop-erties of the double helix because of the presence of a 2-amino group in the minor groove (4), which renders S-2L DNA resistant to most restriction enzymes (5).

The substitution of adenine (A) in S-2L DNA (1) suggested an aminoadenine biosynthetic pathway encoded by the phage. This is in line with the fact that the S-2L genome encodes a putative homolog of succinoadenylate synthase PurA, which catalyzes the first step of de novo biosynthesis of adenosine 5′-monophosphate (AMP) by coupling the hydrolysis of guanosine
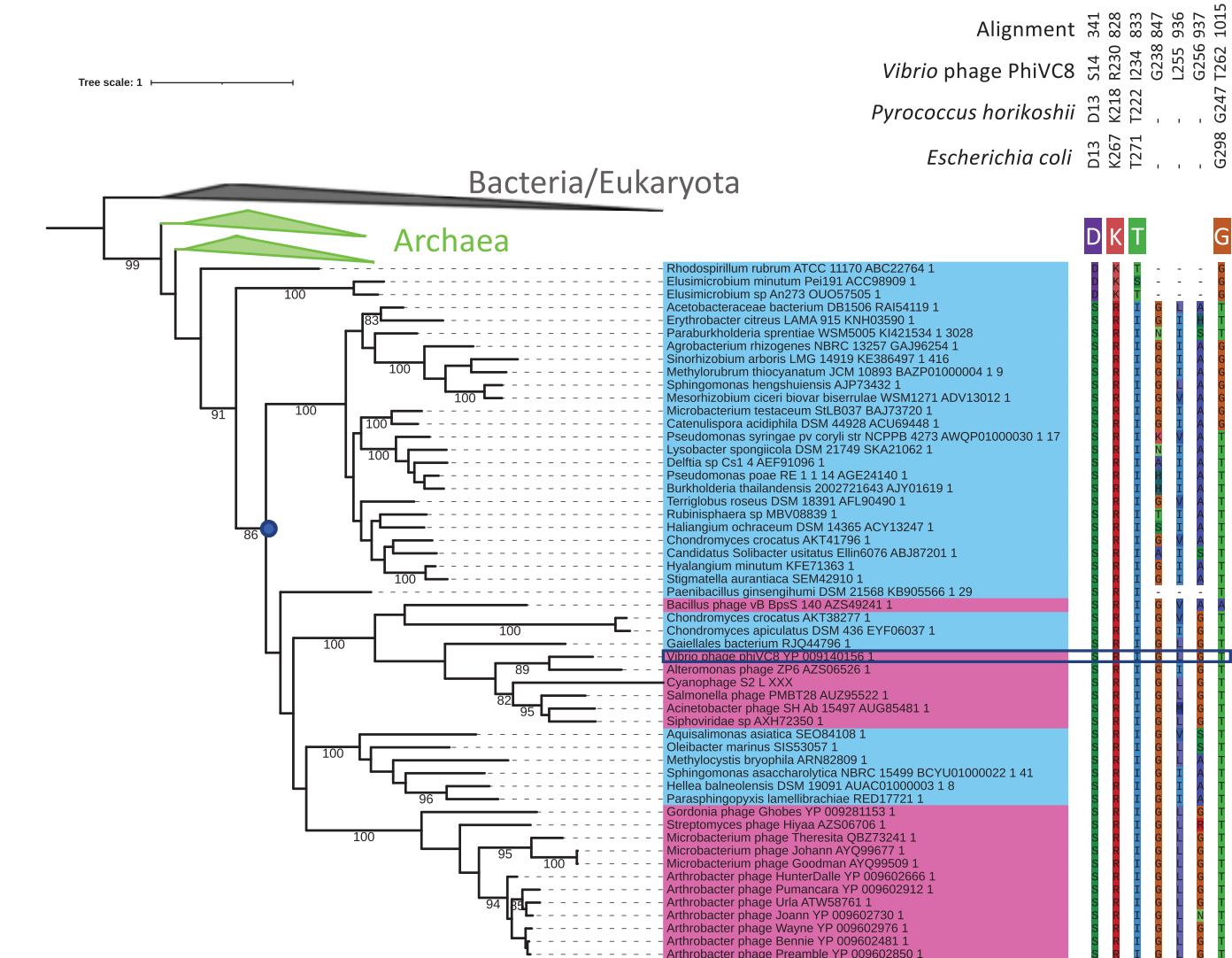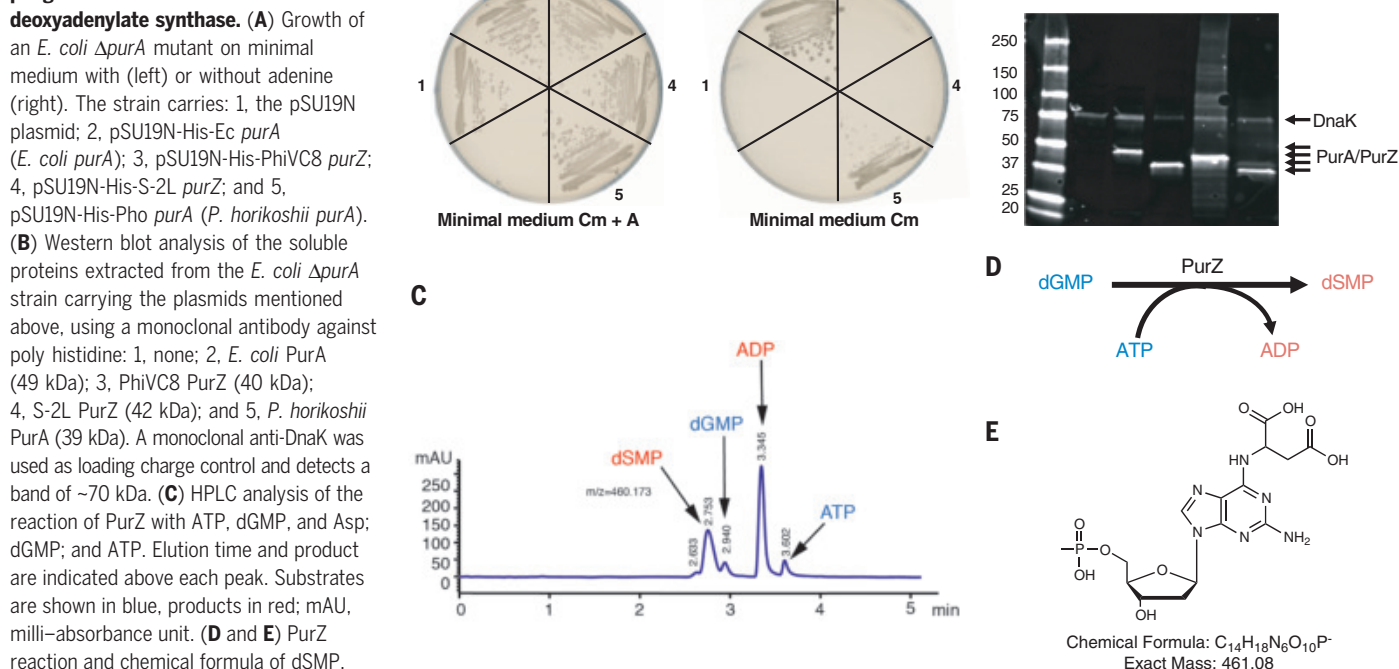
Fig. 1. Phylogenetic placement of PurZ sequences and their specific residues. Maximum likelihood tree of PurA and PurZ homologs. The tree was calculated by RaXML, with the LG+I+G4 model, 506 sequences (including 19 representative phage sequences), and 213 unambiguously aligned amino acid positions. The corresponding full tree and the nontrimmed alignment are provided in fig. S1. The scale bar indicates the average number of substitutions per site. For readability, only bootstrap values >80% are shown. The tree is rooted in between bacterial eukaryal (collapsed) and archaeal (in green, collapsed) canonical PurA sequences. The PurZ clade contains phage (pink) and bacterial sequences (blue). The *Vibrio* phage PhiVC8 PurZ sequence is framed. The seven columns on the right correspond to amino acid residues specific to the PurZ clade (the upper large rectangles indicate the major amino acids of archaeal sequences). The relative positions in PurA from *P. horikoshii*, *E. coli*, and *Vibrio* phage PhiVC8 are indicated. The blue dot indicates the branch where these seven specific residues emerged from canonical PurA. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

5′-triphosphate (GTP) with the synthesis of succinoadenylate from L-aspartate and inosine 5′-monophosphate (IMP) (*6*). We identified PurZ homologs in several bacteriophages, notably in the PhiVC8 phage infecting *Vibrio cholerae* (*7*), whose DNA contains aminoadenine instead of adenine (*8*). This prompted us to characterize the activity of the PurZ enzymes, to elucidate the biosynthetic pathway for aminoadenine nucleotides, and to probe it in vivo.

We first expanded our search for PurA and PurZ homologs and found candidates in 60 other phages infecting distantly related bacteria (mostly *Actinobacteria*, *Firmicutes*, and *Proteobacteria*) (table S1). Phylogenetic analysis revealed a clear distinction between prokaryotic and eukaryotic canonical PurA sequences on the one hand and phage PurZ sequences (pink) falling into a clade embedded within archaeal PurA sequences on the other hand (green) (Fig. 1 and fig. S1). The PurZ clade also includes some bacterial homologs (blue), which could originate from phages. This evolutionary closeness is consistent with two specific deletions that are not found in bacterial and eukaryotic PurA (fig. S1A).

Fig. 2. The PurZ protein of *Vibrio* phage PhiVC8 is a succino-amino-deoxyadenylate synthase. (**A**) Growth of an *E. coli* Δ*purA* mutant on minimal medium with (left) or without adenine (right). The strain carries: 1, the pSU19N plasmid; 2, pSU19N-His-Ec *purA* (*E. coli purA*); 3, pSU19N-His-PhiVC8 *purZ*; 4, pSU19N-His-S-2L *purZ*; and 5, pSU19N-His-Pho *purA* (*P. horikoshii purA*). (**B**) Western blot analysis of the soluble proteins extracted from the *E. coli* Δ*purA* strain carrying the plasmids mentioned above, using a monoclonal antibody against poly histidine: 1, none; 2, *E. coli* PurA (49 kDa); 3, PhiVC8 PurZ (40 kDa); 4, S-2L PurZ (42 kDa); and 5, *P. horikoshii* PurA (39 kDa). A monoclonal anti-DnaK was used as loading charge control and detects a band of ~70 kDa. (**C**) HPLC analysis of the reaction of PurZ with ATP, dGMP, and Asp; dGMP; and ATP. Elution time and product are indicated above each peak. Substrates are shown in blue, products in red; mAU, milli–absorbance unit. (**D** and **E**) PurZ reaction and chemical formula of dSMP.

Fig. 3. Crystal structures of *Vibrio* phage PhiVC8 PurZ.
(**A**) Structure of PurZ without and with various ligands. The region in pink represents the missing loop in case of no ligand or ATP only. (**B**) Fixation of ATP in PurZ (6FM0), hydrogen bonds are indicated. (**C**) Fixation of dGMP in PurZ (6FM1), hydrogen bonds are indicated. (**D** and **E**) Fixation of aspartate in PurZ (6TNH). (E) Loop movement with aspartate binding. (Left) Aspartate interactions with protein backbone. (Right) Superimposition of both structures (6TNH in blue and 6FM1 in orange). Distances from aspartate to residues are indicated. (**F**) PurZ-specific residues are close to the catalytic site (6TNH).
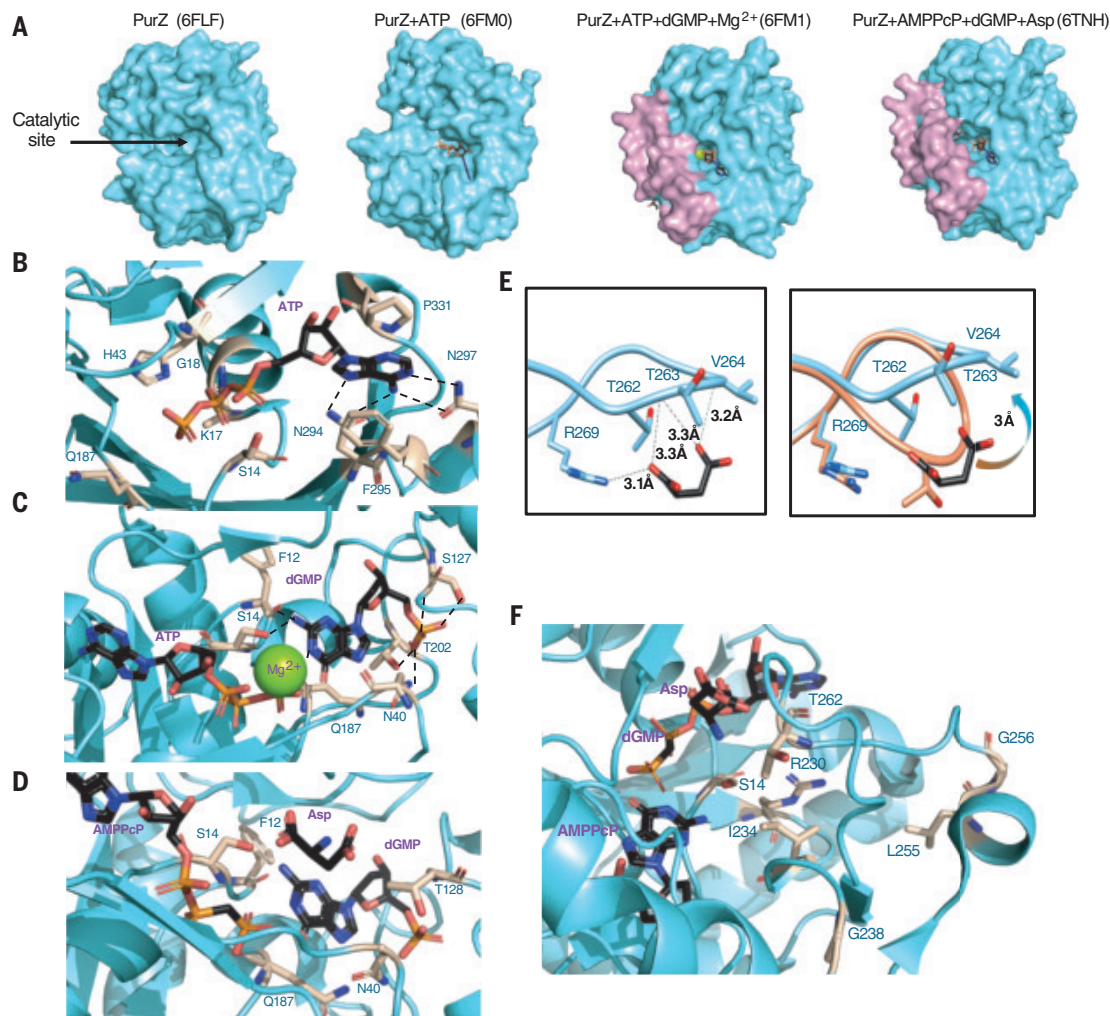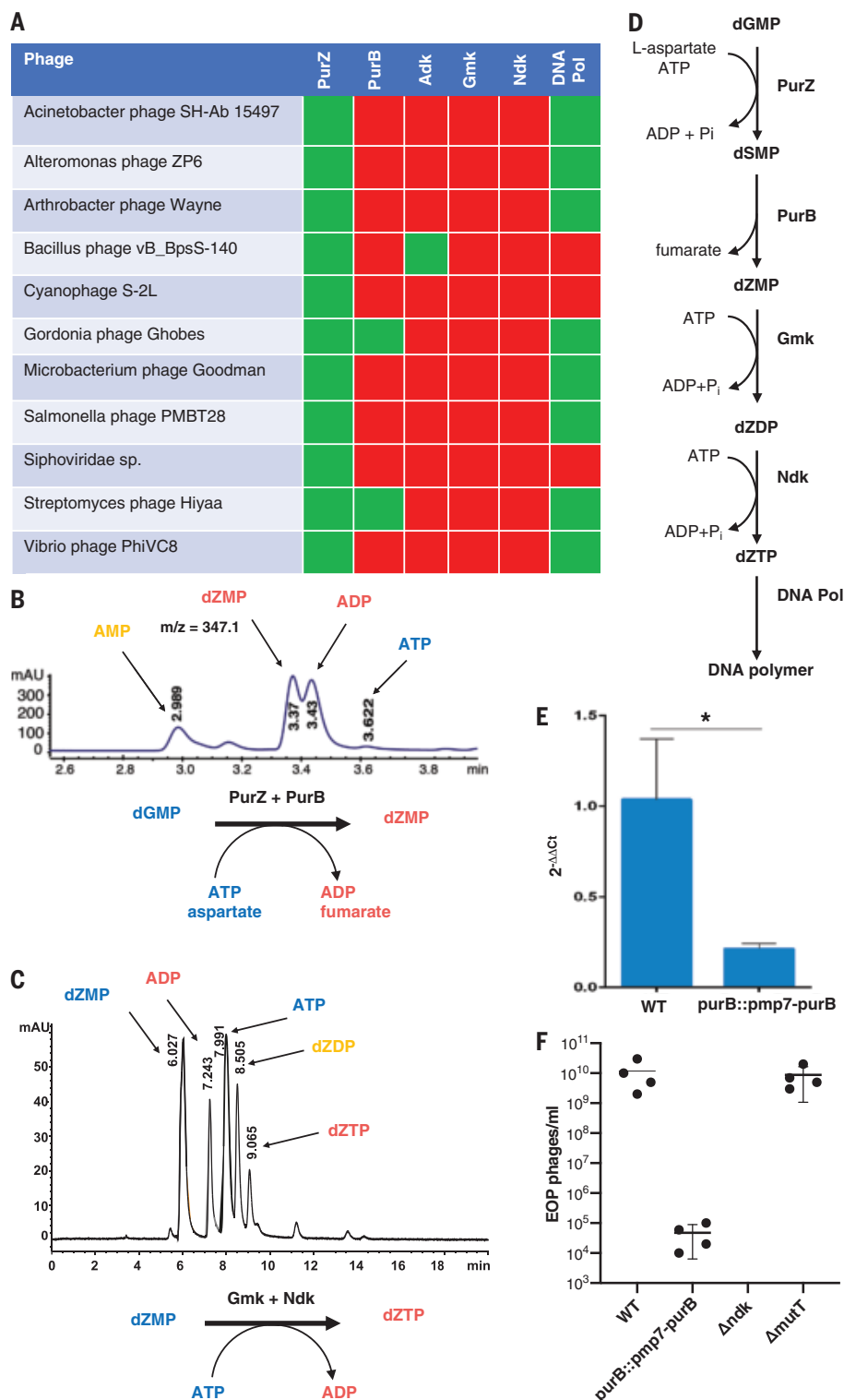
Fig. 4. Proposed pathway for dZTP synthesis. (**A**) Taxonomic distribution of potential dZTP synthesis enzymes for 11 representative phages that have PurZ coding genes. Green squares indicate presence, and the red squares indicate absence. (**B** and **C**) HPLC analysis of substrates and products after enzymatic reactions. The substrates are shown in blue, the products in red, and the intermediates products in yellow. The resulting reaction is displayed on the bottom. (**D**) dZTP synthesis pathway. PurZ, N6-succino-2-amino-2′-deoxyadenylate synthase; PurB, adenylosuccinate lyase; Gmk, guanylate kinase; Ndk, nucleoside diphosphate kinase; DNA Pol, DNA polymerase I. (**E**) Expression of *purB* in the wild type (WT) and purB* mutant (*purB::pmp7-purB*). (**F**) Efficiency of plating (EOP) of different mutants. *Vibrio* phage PhiVC8 is not able to replicate in *V. cholerae* mutants of *purB** and *ndk* in contrast to WT and *mutT*.

The evolutionary distinction between the PurZ clade and cellular PurA is also confirmed by a number of specific sequence signatures corresponding to *Vibrio* phage PhiVC8 amino acids S14, R230, I234, G238, L255, G256, and T262. In particular, the essential D13 residue (S14 in PhiVC8) in *Escherichia coli* PurA (*6*) is not conserved in PurZ sequences (Fig. 1).

We investigated the function of phage PurZ enzymes in a complementation assay using an *E. coli purA* deletant. *E. coli* and *Pyrococcus horikoshii purA* genes restored adenine prototrophy on minimal medium (*9*), whereas no growth was observed with PhiVC8 and S-2L *purZ* genes (Fig. 2A), although expressed adequately in *E. coli* (Fig. 2B). We next characterized the enzymatic activity of PurZ by rapid-resolution high-performance liquid chromatography (RR-HPLC). No standard reaction was catalyzed by PhiVC8 or S-2L PurZ using GTP, IMP, L-aspartate, and Mg^{2+}, whereas a reaction occurred for both PurZs only in the presence of deoxyguanosine monophosphate (dGMP), L-aspartate, adenosine 5′-triphosphate (ATP), and Mg^{2+}. The reaction produced a peak in addition to adenosine 5′-diphosphate (ADP) (retention time, 2.753 min) (Fig. 2C), which was confirmed by nuclear magnetic resonance (NMR) as N6-succino-2-amino-2′-deoxyadenylate (dSMP)—a nucleotide not previously described (Fig. 2D and fig. S2).

We solved the crystal structure of apo-PurZ at 1.3-Å resolution [Protein Data Bank identifier (PDB ID): 6FLF] by molecular replacement using the coordinates of *P. horikoshii* OT3 PurA (PDB ID: 2D7U) as a template (fig. S3A). We next screened potential ligands by crystallography (fig. S3A) and obtained the structures of PhiVC8 PurZ in complex with ATP (resolution = 1.7 Å; PDB ID: 6FM0); with dGMP, ATP, and Mg^{2+} (2.3 Å; PDB ID: 6FM1); and with β,γ-methyleneadenosine 5′-triphosphate (AMPPcP), dGMP, and aspartate (2.3 Å; PDB ID: 6TNH).

The secondary structure and the overall polypeptide fold of PhiVC8 PurZ exhibit a typical Rossmann fold with a central β sheet of nine β strands flanked by four α helices on each side (fig. S3B). In apo-PurZ and PurZ-ATP structures, three loops (G110-S129, G234-G238, and E246-V262) were disordered and could not be fitted in the electron density map (Fig. 3A). Cocrystallization of PurZ with ATP, dGMP, and Mg^{2+} or with AMPPcP, dGMP, and aspartate allowed us to solve the structure of these regions, which suggests that fixation of ATP or AMPPcP and dGMP induces the stabilization of this part of the protein (Fig. 3A).

In the structure of PurZ-ATP (6FM0), the adenine ring of ATP is stabilized by stacking interactions between F295 and P331, and it interacts with the side chains of N294 and N297 through hydrogen bonds (Fig. 3B). The

ribose moiety does not make any direct interaction with the residue in the enzyme active site of the enzyme, although the α phosphate interacts with G18, the β phosphate interacts with the backbone and the side chains of both H43 and K17, and the γ phosphate interacts with the backbone of Q187, G18, and S14 and with the side chain of K17.

In the crystal structure of PurZ obtained in the presence of ATP, dGMP, and Mg^{2+} (6FM1), the density corresponding to dGMP and a Mg^{2+} cation in a typical octahedral coordination was observed. The guanine heterocycle interacts through two hydrogen bonds with the protein backbone (Q187 and F12) and with the hydroxyl group of S14 (Fig. 3C). The ribose moiety does not interact directly with the protein backbone. The phosphate group of dGMP is within hydrogen-bonding distance (between 2.9 and 3.2 Å) from T202, N40, S127, and the R142 of the monomer B in the asymmetric unit. The β and γ phosphates of ATP coordinate the Mg^{2+} cation present in the active site (Fig. 3C).

The structure of PhiVC8 PurZ in complex with dGMP, L-aspartate, and AMPPcP, a non-hydrolyzable analog of ATP, showed electron densities corresponding to the different ligands in one of the two monomers present in the asymmetric unit (6TNH). The aspartate substrate interacts mainly with the protein backbone (T262, T263, V264, and R269) and is ideally positioned to react with the dGMP molecule (Fig. 3, D and E).

To further investigate the determinants of PurZ specificity, we mapped the seven residues specific for the PurZ clade (Fig. 1) in the structure. These residues are all localized near the catalytic site (Fig. 3F), S14 and I234 lying at the center of the catalytic site and at a distance compatible with a hydrogen bond with the amine group of the C2 of dGMP. The side chain of R230 (K267 in *E. coli*) is located near the backbone of S14 (3.0 to 3.8 Å). T262 (G298 in *E. coli*) is in the loop that moves when aspartate is bound to PurZ, which suggests a role of this residue during catalysis. Finally, G256, L255, and G238 belonging to an unstructured region facing the catalytic pocket may contribute to the enzymatic reaction (Fig. 3A).

Additional enzymes are required to convert dSMP to 2-amino-2′-deoxyadenosine-5′-triphosphate (dZTP). Because PurZ-encoding genomes do not have any lyase or kinase (Fig. 4A), we hypothesized that the host enzymes were recruited. Adenylosuccinate lyase PurB from *V. cholerae* converted dSMP to dZMP in the presence of PhiVC8 PurZ, ATP, L-aspartate, and Mg^{2+} (Fig. 4B and fig. S4, A to C). dZMP was in turn converted into dZTP using ATP by both guanylate and nucleoside diphosphate kinases (Gmk and Ndk) (Fig. 4, C and D). Altogether, this led us to propose the dZTP pathway (Fig. 4D).

PurZ is essential for phage proliferation, whereas *gmk* and *amk* are essential for *V. cholerae* O1 viability, so we constructed a *V. cholerae* O1 mutant *purB** with reduced expression of *purB* (Fig. 4E) and then generated a *ndk* deletant as well as a *mutT* (as a control). The *mutT* deletant is comparable to the wild type, the *purB** mutant showed a 10^{-5} reduction in PhiVC8 infection efficiency, and no infection was observed in the *ndk* mutant (Fig. 4F). These results suggest that the host genes are essential for the production of phage progeny and are needed for dZTP biosynthesis. Introduction of the dZTP pathway into living cells is expected to expand the range of synthetic and analytical options in nanotechnology as well as increase the understanding of the adaptation to chemical changes of the genetic material.

REFERENCES AND NOTES

1. M. D. Kirnos, I. Y. Khudyakov, N. I. Alexandrushkina, B. F. Vanyushin, *Nature* **270**, 369–370 (1977).
2. M. P. Callahan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13995–13998 (2011).
3. I. Y. Khudyakov, M. D. Kirnos, N. I. Alexandrushkina, B. F. Vanyushin, *Virology* **88**, 8–18 (1978).
4. C. Bailly, M. J. Waring, *Nucleic Acids Res.* **26**, 4309–4314 (1998).
5. M. Szekeres, A. V. Matveyev, *FEBS Lett.* **222**, 89–94 (1987).
6. R. B. Honzatko, H. J. Fromm, *Arch. Biochem. Biophys.* **370**, 1–8 (1999).
7. A. Solís-Sánchez *et al.*, *Virol. J.* **13**, 47 (2016).
8. V. Pezo *et al.*, *Science* **372**, 520–524 (2021).
9. A. Bouyoub, G. Barbier, P. Forterre, B. Labedan, *J. Mol. Biol.* **261**, 144–154 (1996).
10. P. Garcia, Sleiman_et_al_Sup_data, dataset, Figshare (2021); https://doi.org/10.6084/m9.figshare.14134922.v1.

NONCANONICAL GENOME

Noncanonical DNA polymerization by aminoadenine-based siphoviruses

Valerie Pezo[1], Faten Jaziri[1]†, Pierre-Yves Bourguignon[2,3], Dominique Louis[4], Deborah Jacobs-Sera[5], Jef Rozenski[6], Sylvie Pochet[7], Piet Herdewijn[6], Graham F. Hatfull[5], Pierre-Alexandre Kaminski[8], Philippe Marliere[1,3]*

Bacteriophage genomes harbor the broadest chemical diversity of nucleobases across all life forms. Certain DNA viruses that infect hosts as diverse as cyanobacteria, proteobacteria, and actinobacteria exhibit wholesale substitution of aminoadenine for adenine, thereby forming three hydrogen bonds with thymine and violating Watson-Crick pairing rules. Aminoadenine-encoded DNA polymerases, homologous to the Klenow fragment of bacterial DNA polymerase I that includes 3′-exonuclease but lacks 5′-exonuclease, were found to preferentially select for aminoadenine instead of adenine in deoxynucleoside triphosphate incorporation templated by thymine. Polymerase genes occur in synteny with genes for a biosynthesis enzyme that produces aminoadenine deoxynucleotides in a wide array of *Siphoviridae* bacteriophages. Congruent phylogenetic clustering of the polymerases and biosynthesis enzymes suggests that aminoadenine has propagated in DNA alongside adenine since archaic stages of evolution.

The nitrogenous base 2-aminoadenine (also known as 2,6-diaminopurine, hereafter abbreviated as Z) embodies an extreme chemical deviation reached during the evolution of the biosphere. Its molecular formula $C_5H_6N_6$ has the most biased compositional ratio (N/C = 1.2) known among natural nucleobases. It completely replaces adenine in the DNA of the lytic phage S-2L, a siphovirus that preys on freshwater cyanobacteria of the genus *Synechococcus* (*1*).

This phage, whose double-stranded linear DNA contains about 35% G, 35% C, 15% T and 15% Z, was discovered in Vanyushin's laboratory

in 1977 (2). Like the G:C pair, the Z:T pair is stabilized by three hydrogen bonds (fig. S1), as opposed to the two in A:T (3), thus singularly violating canonical Watson-Crick pairing rules (4).

The S-2L genome contains a gene, designated *purZ*, that encodes an enzyme distantly related to succinoadenylate synthase (EC 6.3.4.4, encoded by *purA*) in the adenine biosynthetic pathway (5). PurZ catalyzes the first step of a metabolic pathway that converts deoxyguanosine monophosphate (dGMP) into 2-amino-2′-deoxyadenosine 5′-triphosphate (dZTP), a putative DNA polymerase substrate (Fig. 1A). However, a gene encoding DNA polymerase was not detected in the S-2L genome (GenBank AX955019.1). Detailed characterization of the PurZ enzyme and the pathway from dGMP to dZTP is the focus of a companion article (6).

To find the missing DNA polymerase that uses dZTP as a substrate, we first searched protein databases (UniProt), using the PurZ amino acid sequence encoded by S-2L as our query, and identified a cluster of homologs that all belong to the *Siphoviridae* bacteriophages. Such siphoviruses infect hosts as distant as proteobacteria (e.g., *Vibrio*, *Salmonella*, and *Acinetobacter*), actinobacteria (e.g., *Arthrobacter* and *Gordonia*), and cyanobacteria (e.g., *Synechococcus*), which dwell in habitats as diverse as soil, fresh water, and seawater (table S1).

Notably, a *polA* gene homolog was found to occur in synteny with *purZ* in all of these phage genomes, but not in that of S-2L (Fig. 1B). The PolA homologs, which we called DpoZ, corresponded to the Klenow fragment of *Escherichia coli* DNA polymerase I, lacking the 5′-exonuclease domain but retaining the 3′-exonuclease domain (figs. S2 and S3). A list of phage genomes that presumably contain aminoadenine can be found in table S1. The conserved synteny between the *dpoZ* and *purZ* genes in siphoviral genomes (Fig. 1B) suggests that the two enzymatic functions coexisted in a common ancestor encoded with aminoadenine and have since coevolved.

To test this hypothesis, we reconstructed phylogenetic trees separately for the PurZ and DpoZ amino acid sequences by strictly applying the same algorithmic method [multiple alignment of sequence families followed by a posteriori average reconstruction using a general time reversible model (7)]. Two almost perfectly congruent unrooted trees were thus obtained (Fig. 1C). They deviate in the slightly different placement of the *Streptomyces* Hiyaa phage, whose genome is twice as large [83 kilo–base pairs (kbp)] as those of other siphoviruses (table S1).

The S-2L phage is unusual because its genome has the lowest Z:T content among phages that encode a *purZ* gene and because it lacks a *dpoZ* gene (table S1). The PurZ tree places S-2L in a branch with phages that infect proteobacteria, which are endowed with a Gram-negative double-membrane apparatus like the cyanobacterial host of S-2L, whereas actinobacterial hosts of other siphoviruses are Gram-positive (Fig. 1C).

The ancient origin of aminoadenine is strongly suggested by the deep and congruent phylogeny of PurZ and DpoZ (Fig. 1C), which indicates that it has been used as an information carrier among siphoviruses since the evolutionary divergence of actinobacteria, cyanobacteria, and proteobacteria ~3.5 billion years ago (8). An important question for further research is whether the use of dZTP emerged as a shift of function [i.e., an exaptation (9)] of the cellular enzymes PurA and PolA, which were recruited and repurposed by a siphoviral ancestor, or whether it paralleled the use of 2′-deoxyadenosine 5′-triphosphate (dATP) since earlier stages of evolution.

To ascertain the base composition of the siphoviral genomes containing *purZ* and *dpoZ*, we first showed that the restriction enzymes that cleave adenine-containing hexamers did not digest DNA from the *Vibrio* phage (PhiVC8), *Gordonia* phage (Ghobes), and *Arthrobacter* phage (Wayne), as would have been expected if they contained aminoadenine (Fig. 2A). Chemical evidence for the absence of deoxyadenosine (dA) and presence of 2-aminodeoxyadenosine (dZ) was provided by high-performance liquid chromatography (HPLC) fractionation and mass spectrometry applied to enzymatic hydrolysates of phage DNA (Fig. 2B). A minor fraction of dA was found together with dZ in the DNA from the *Gordonia* phage Ghobes (Fig. 2B).

We then investigated whether DpoZ discriminated between aminoadenine and adenine in replication. As shown in Fig. 3B, dZTP led to the formation of full-length products in response to the $(dT)_{24}$ template, using the His-tagged versions of the four phage polymerases produced in *E. coli* and purified to homogeneity. By contrast, dATP led to incomplete polymerization products under the same conditions. The preference for Z over

A was also observed in the template. Polymerization of dTTP (3′-deoxythymidine 5′-triphosphate) catalyzed by DpoZ enzymes was more efficient in response to template $(dZ)_{24}$ than to $(dA)_{24}$ (Fig. 3B). Notably, the Klenow fragment of *E. coli* PolA (His-tagged in the same way) catalyzed the polymerization of full-length product with all combinations of substrates and templates, showing little discrimination for or against Z (Fig. 3B).

As shown in Table 1, quantitative confirmation of the discrimination effect qualitatively displayed in Fig. 3B was obtained by measuring the affinity (Michaelis constant, K_M) and turnover number (k_{cat}) of the DpoZ enzymes and comparing with the values for Klenow PolA. The DpoZ enzymes from the *Vibrio* and *Acinetobacter* phages were, respectively, 90 and 29 times as efficient in using dZTP instead of dATP as a substrate, whereas the Klenow enzyme slightly favored dATP. These results were obtained by using versions of the enzymes that carry a mutation that disables the active site of the 3′-exonuclease domain, so as to univocally test polymerase activity (figs. S4 and S5).

To study polymerization in the context of the natural history of phages, we also performed primer extension assays by using templates corresponding to a 50-nucleotide sequence from the siphoviral genome SH-Ab 15497 (Fig. 3, C to E). Judging by the size of the elongation products, the dZTP substrate was preferred instead of dATP by siphoviral DNA polymerases, regardless of whether the template contained Z or A (Fig. 3, D and E). This result is in contrast to findings for Klenow PolA, which showed a preference for dATP and the A-containing template.

The exclusive incorporation in DNA of a noncanonical base such as aminoadenine may result from multiple enzymatic activities encoded by a phage genome (10). At the prereplicative stage, the canonical precursor dATP may be cleaved by a phosphohydrolase that spares dZTP. At the replicative stage, the incorporation of dZTP may be favored over dATP by polymerase activity. At the postreplicative stage, excision of incorporated dA may be promoted relative to that of dZ by 3′-exonuclease activity.

The main finding of this study is the identification of a category of DNA polymerases that are encoded with a noncanonical nucleobase and that discriminate against the incorporation of the canonical counterpart at the replicative stage. In short, the biochemical

[1]Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ. Evry, Université Paris-Saclay, 2 Rue Gaston Crémieux, 91057 Evry, France. [2]Werkstatt fuer Potenzielle Genetik, Naunynstrasse 30, 10997 Berlin, Germany. [3]TESSSI, 81 Rue Réaumur, 75002 Paris, France. [4]Alderys, 86 Rue de Paris, 91400 Orsay, France. [5]Department of Biological Sciences, University of Pittsburgh, 4249 Fifth Avenue, Pittsburgh, PA 15260 USA. [6]Laboratory of Medicinal Chemistry, Rega Institute for Biomedical Research, KU Leuven, Herestraat 49, Box 1041, 3000 Leuven, Belgium. [7]Organic Chemistry, CNRS UMR3523, Department of Chemistry and Biocatalysis, Institut Pasteur, 25-28 Rue du Docteur Roux, 75015 Paris, France. [8]Biology of Gram-Positive Pathogens, CNRS URL3526, Department of Microbiology, Institut Pasteur, 25-28 Rue du Docteur Roux, 75015 Paris, France.
†Present address: DNA Script, 67 Avenue de Fontainebleau, 94270 Le Kremlin-Bicêtre, France.
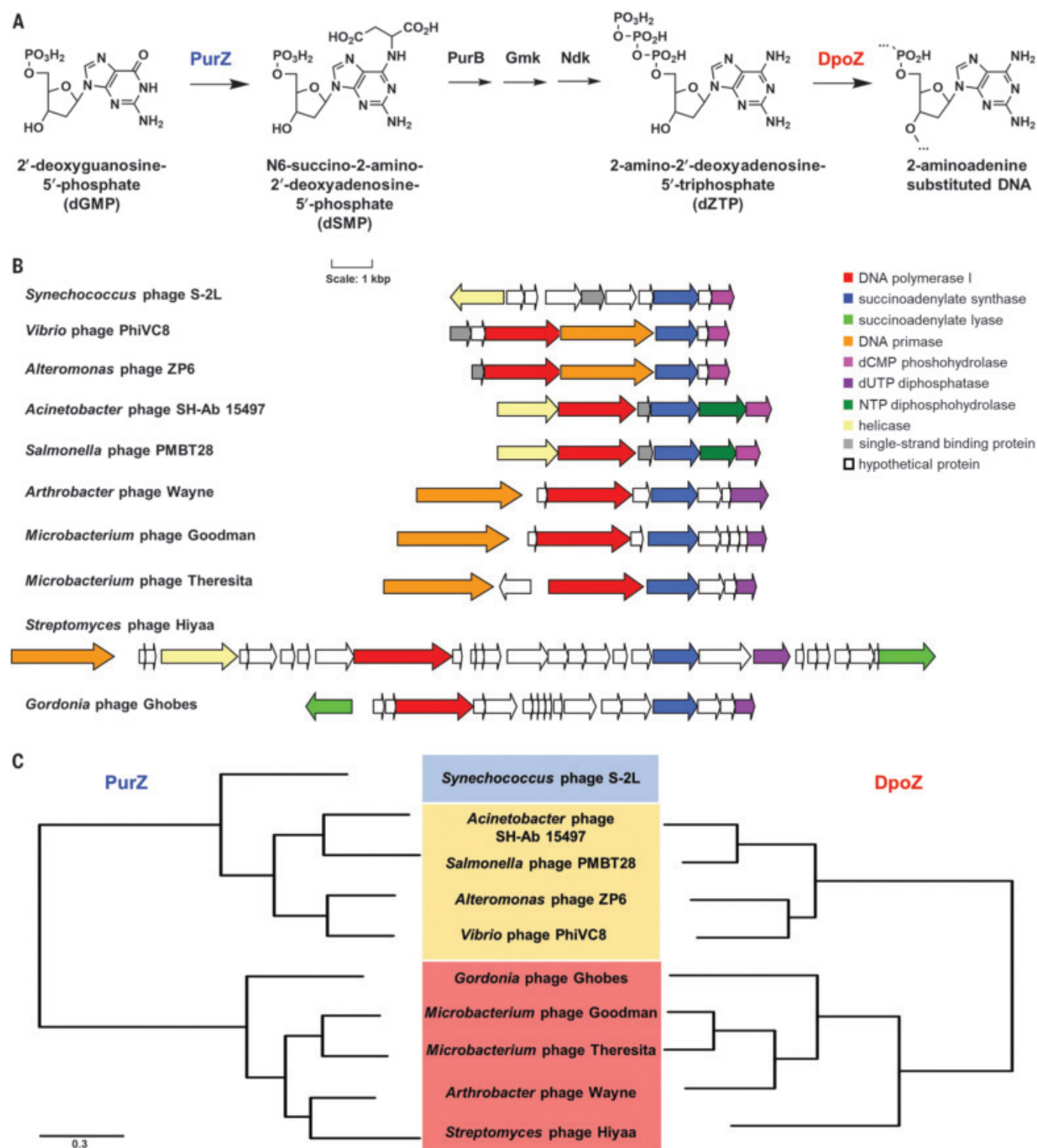*Corresponding author. Email: phmarliere@tesssi.eu

Fig. 1. Genetic synteny and molecular phylogeny of dZTP-cognate enzymes from siphoviruses. (**A**) Incorporation pathway of dZTP via reactions catalyzed by the phage-encoded biosynthetic enzyme PurZ and DNA polymerase DpoZ. The steps labeled "PurB," "Gmk," and "Ndk" are catalyzed by enzymes of the bacterial host (*6*). (**B**) Comparative arrangement of genes in siphovirus genomes. The alignment is set on the purZ gene (in blue); a color code is given for putative enzymes inferred from homology with nucleotide-metabolizing enzymes (*19*). dCMP, deoxycytidine 5′-monophos-phate; dUTP, deoxyuridine 5′-triphosphate; NTP, nucleotide triphosphate. (**C**) Mirror representation of phylogenetic trees for the PurZ and DpoZ families. Each tree was reconstructed separately using the phylogenetically informative sites of the multiple sequence alignment for the corresponding family (fig. S6). Branch distances in both trees follow the same scale and correspond to the expected number of mutations per aligned position. Phages are grouped according to the phylum of their bacterial hosts: cyan, cyanobacteria; yellow, proteobacteria; magenta, actinobacteria.
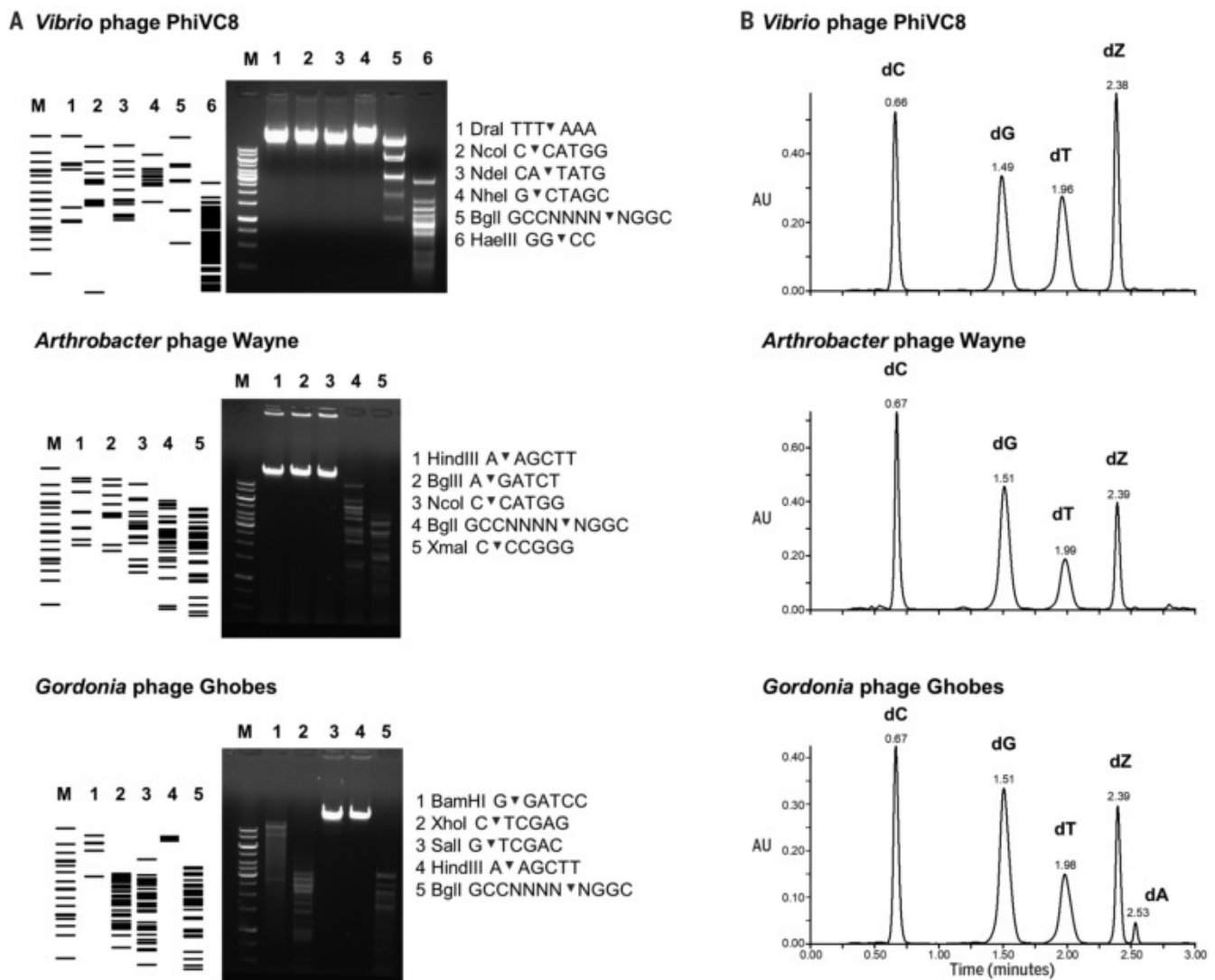
A *Vibrio* phage PhiVC8

1 DraI TTT▼AAA
2 NcoI C▼CATGG
3 NdeI CA▼TATG
4 NheI G▼CTAGC
5 BglI GCCNNNN▼NGGC
6 HaeIII GG▼CC

Arthrobacter phage Wayne

1 HindIII A▼AGCTT
2 BglII A▼GATCT
3 NcoI C▼CATGG
4 BglI GCCNNNN▼NGGC
5 XmaI C▼CCGGG

Gordonia phage Ghobes

1 BamHI G▼GATCC
2 XhoI C▼TCGAG
3 SalI G▼TCGAC
4 HindIII A▼AGCTT
5 BglI GCCNNNN▼NGGC

B *Vibrio* phage PhiVC8

Arthrobacter phage Wayne

Gordonia phage Ghobes

Fig. 2. Detection of 2-amino-2′-deoxyadenosine (dZ) in bacteriophage DNA. (**A**) Pattern of restriction endonuclease cleavage. Predicted profiles are shown on the left and observed profiles on the right. With the exception of BamHI and XhoI (*Gordonia* phage Ghobes), endonucleases specific for adenine-containing restriction sites (DraI, NcoI, NdeI, NheI, HindIII, BglII, and SalI) did not cut DNA from the three phages. M, DNA ladder. (**B**) HPLC profiles of phage DNA digested by nuclease and phosphatase. AU, arbitrary units; dC, deoxycytidine; dG, deoxyguanosine; dT, deoxythymidine; dZ, deoxyaminoadenosine; dA, deoxyadenosine.

Table 1. Enzymology of the discrimination between dZTP and dATP. The average and standard deviation (±SD) of three independent assays are given for each value of K_M and k_{cat}. His-tagged versions of 3′-exonuclease mutant DpoZ polymerases from phages are compared with the His-tagged 3′-exonuclease mutant of *E. coli* PolA (Klenow fragment).

DNA polymerase	K_M (µM)		$\dfrac{K_M^{(A)}}{K_M^{(Z)}}$	k_{cat} (s^{-1})		$\dfrac{k_{cat}^{(Z)}}{k_{cat}^{(A)}}$	Efficiency (µM^{-1} s^{-1})		$\dfrac{K_M^{(A)} k_{cat}^{(Z)}}{K_M^{(Z)} k_{cat}^{(A)}}$
	dZTP	dATP		dZTP	dATP		dZTP	dATP	
E. coli Klenow fragment	0.81 ± 0.14	0.70 ± 0.17	0.86	13.96 ± 1.99	20.61 ± 1.52	0.68	17.73 ± 4.96	30.40 ± 5.82	0.58
Vibrio phage PhiVC8	0.10 ± 0.02	4.50 ± 1.38	45.00	0.42 ± 0.14	0.21 ± 0.05	2.00	4.35 ± 1.56	0.048 ± 0.017	90.62
Acinetobacter phage SH-Ab 15497	1.30 ± 0.36	21.40 ± 3.40	16.47	0.29 ± 0.07	0.16 ± 0.01	1.81	0.23 ± 0.06	0.008 ± 0.001	28.75

Fig. 3. Discrimination between aminoadenine and adenine by DNA polymerases. Primer extension assays were performed using purified His-tagged polymerases, phage DpoZ, and *E. coli* PolA Klenow fragment. (**A**) Experimental setup showing nucleotide sequences of duplexes between a fluorescent (FAM)–labeled primer (X1903) and a homopolymer template [X1904, (dT)$_{24}$; X1930, (dA)$_{24}$; X2009, (dZ)$_{24}$]. (**B**) Elongation products polymerized for 30 min at 37°C were loaded on denaturing 17% polyacrylamide gels. Lane numbers refer to conditions indicated in (A). (**C**) Experimental setup using duplexes between a fluorescent primer and a heteropolymer template corresponding to a 50-nucleotide sequence from the SH-Ab 15497 phage genome. (**D**) Elongation products polymerized from a Z-containing template (X2586) and primer (X2587). (**E**) Elongation products polymerized from an A-containing template (X2364) and primer (X2365).

synthesis and replication of Z, as well as the exclusion of A, are encrypted in the genetic text written with Z. Such a natural example of "molecular ostracism" bodes well for the field of synthetic biology, as it provides a proof of principle for chemically diversifying nucleic acids in vivo (*11–18*).

REFERENCES AND NOTES

1. I. Y. Khudyakov, M. D. Kirnos, N. I. Alexandrushkina, B. F. Vanyushin, *Virology* **88**, 8–18 (1978).
2. M. D. Kirnos, I. Y. Khudyakov, N. I. Alexandrushkina, B. F. Vanyushin, *Nature* **270**, 369–370 (1977).
3. C. Bailly, M. J. Waring, *Nucleic Acids Res.* **26**, 4309–4314 (1998).
4. J. D. Watson, F. H. Crick, *Nature* **171**, 964–967 (1953).
5. P. Marlière et al., Patent WO03093461 (2003).
6. D. Sleiman et al., *Science* **372**, 516–520 (2021).
7. P. Benner, M. Bačák, P. Y. Bourguignon, *Bioinformatics* **30**, i534–i540 (2014).
8. P. P. Sheridan, K. H. Freeman, J. E. Brenchley, *Geomicrobiol. J.* **20**, 1–14 (2003).
9. S. J. Gould, E. S. Vrba, *Paleobiology* **8**, 4–15 (1982).
10. P. Weigele, E. A. Raleigh, *Chem. Rev.* **116**, 12655–12687 (2016).
11. A. Rich, in *Horizons in Biochemistry*, M. Kasha, B. Pullmann, Eds. (Academic, 1962), pp. 103–126.
12. P. Herdewijn, P. Marlière, *Chem. Biodivers.* **6**, 791–808 (2009).
13. A. Ravikumar, A. Arrieta, C. C. Liu, *Nat. Chem. Biol.* **10**, 175–177 (2014).
14. S. Hoshika et al., *Science* **363**, 884–887 (2019).
15. D. A. Malyshev et al., *Nature* **509**, 385–388 (2014).
16. P. Marlière et al., *Angew. Chem. Int. Ed.* **50**, 7109–7114 (2011).
17. A. P. Mehta et al., *J. Am. Chem. Soc.* **138**, 14230–14233 (2016).
18. M. Luo et al., *J. Am. Chem. Soc.* **141**, 10844–10851 (2019).
19. K. J. Harrison, V. de Crécy-Lagard, R. Zallot, *Bioinformatics* **34**, 1406–1408 (2018).

CORONAVIRUS

Structural impact on SARS-CoV-2 spike protein by D614G substitution

Jun Zhang[1,2]†, Yongfei Cai[1,2]†, Tianshu Xiao[1,2], Jianming Lu[3], Hanqin Peng[1], Sarah M. Sterling[4,5], Richard M. Walsh Jr.[4,5], Sophia Rits-Volloch[1], Haisun Zhu[6], Alec N. Woosley[6], Wei Yang[6], Piotr Sliz[1,2,5], Bing Chen[1,2]*

Substitution for aspartic acid (D) by glycine (G) at position 614 in the spike (S) protein of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) appears to facilitate rapid viral spread. The G614 strain and its recent variants are now the dominant circulating forms. Here, we report cryo–electron microscopy structures of a full-length G614 S trimer, which adopts three distinct prefusion conformations that differ primarily by the position of one receptor-binding domain. A loop disordered in the D614 S trimer wedges between domains within a protomer in the G614 spike. This added interaction appears to prevent premature dissociation of the G614 trimer—effectively increasing the number of functional spikes and enhancing infectivity—and to modulate structural rearrangements for membrane fusion. These findings extend our understanding of viral entry and suggest an improved immunogen for vaccine development.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), an enveloped positive-stranded RNA virus, is the cause of the COVID-19 pandemic (1). Although the viral evolution is slowed by the RNA proofreading capability of its replication machinery (2), a variant with a single-residue substitution (D614G) in its spike (S) protein rapidly became the dominant strain throughout the world (3). It has since further evolved to give several variants of concern (VOCs) (4–6). The trimeric S protein decorates the viral surface and is an important target for the development of diagnostics, therapeutics, and vaccines; therefore, understanding the effect of key mutations may guide intervention strategies. Here, we focus on the D614G mutation that is in all currently circulating strains of SARS-CoV-2. The S protein is produced as a single-chain precursor and is subsequently processed by a furin-like protease into the receptor-binding fragment S1 and the fusion fragment S2 (7). After engagement of the receptor-binding domain (RBD) in S1 with the viral receptor angiotensin-converting enzyme 2 (ACE2) on the host cell surface, followed by a second proteolytic cleavage within S2 (S2′ site) (8), the S protein undergoes large conformational changes, which result in the dissociation of S1 and the irreversible refolding of S2 into a postfusion structure (9, 10). This induces fusion of the virus and host cell membranes to initiate infection. Rapid advances in structural biology of the SARS-CoV-2 S protein include structures of its soluble fragments: the ectodomain stabilized in its prefusion conformation (11–13), RBD-ACE2 complexes (14–17), and segments of S2 in the postfusion state (18). In the prefusion ectodomain structure, S1 folds into four domains—the N-terminal domain (NTD), the RBD, and two C-terminal domains (CTDs)—and wraps around the prefusion S2, with the RBD sampling two distinct conformations: up for a receptor-accessible state and down for a receptor-inaccessible state. We and others have reported structures of a purified, full-length D614 S protein in both prefusion and postfusion conformations (19, 20). Studies by cryo–electron tomography (cryo-ET), with chemically inactivated SARS-CoV-2 preparations, using both D614 and G614 variants have revealed additional structural details of S proteins present on the surface of the virion (21–24).

Epidemiological surveillance has indicated that the SARS-CoV-2 carrying G614 outcompeted the original virus and became the globally dominant form within a month (3, 25, 26). This single-residue substitution appears to correlate with high viral loads in infected patients and high infectivity of pseudotyped viruses, but not with disease severity (3). The G614 virus has comparable sensitivity to neutralization by convalescent human sera or vaccinated hamster sera (3, 27–30), which suggests that vaccines containing D614 remain effective against the G614 virus. Moreover, S1 dissociates more readily from the D614 virus than from G614 virus (31), which indicates that the D614 viral spike is less stable than that of the G614 variant. The G614 ectodomain trimer is reported to sample the RBD-up conformations more frequently than the D614 trimer (13, 29, 32), but it is puzzling why the former binds more weakly to recombinant

ACE2 than the latter (32). The known S trimer structures indicate that the D614G change breaks a salt bridge between D614 and a lysine residue (K854) in the fusion peptide–proximal region (FPPR) (19, 33, 34), which may help clamp the RBD in the prefusion conformation. This observation can explain why the G614 trimer favors the RBD-up conformations but does not account for its increased stability. To resolve these issues, we report the structural consequences of the D614G substitution in the context of the full-length S protein.

We compared the membrane fusion activity of the full-length G614 S protein (fig. S1) with that of the D614 S construct in a cell-cell fusion assay (19). All of the cells expressing S fused efficiently with cells transfected with a human ACE2 construct (fig. S2A), demonstrating that the S proteins expressed on the cell surfaces are fully functional. At low transfection levels, G614 S had higher fusion activity than the D614 S, but the difference diminished with the increased amount of transfected DNA, which suggests that the high expression levels can compensate for lower fusion efficiency of the D614 S protein. The G614 trimer remains sensitive to inhibition by an engineered trimeric ACE2-based inhibitor that competes with the receptor on the target cells (35) (fig. S2B). For protein purification, we used a construct fused with a C-terminal strep-tag, which was equally active in cell-cell fusion as the untagged version (fig. S2A), and purified both G614 and D614 proteins under identical conditions. The D614 protein eluted in three peaks characterized previously as the prefusion S trimer, the postfusion S2 trimer, and the dissociated monomeric S1 (19). The G614 protein eluted as a single major peak, corresponding to the prefusion trimer (Fig. 1A). This suggests that D614G has a notable effect on the stability of the SARS-CoV-2 S trimer. Coomassie-stained SDS–polyacrylamide gel electrophoresis (SDS-PAGE) analysis shows that G614 elutes mainly as the prefusion trimer, comprising the cleaved S1-S2 complex (~90%) and a small amount of the uncleaved S precursor (~10%). We next measured binding of the prefusion trimer fractions of the full-length proteins to recombinant soluble ACE2 by biolayer interferometry (BLI) (Fig. 1B). The S trimers bound more strongly to a dimeric ACE2 than to a monomeric ACE2, as expected. The G614 protein bound ACE2 less tightly than the D614 protein, which is consistent with the measurements reported by others using soluble constructs (32). This observation appears inconsistent with accounts that the G614 trimer has a more exposed RBD than the D614 trimer (13, 21, 22, 29, 32). We note that the second binding event between dimeric ACE2 and a G614 trimer has both a slower on-rate and a slower off-rate than that for a D614 trimer (table S1).

[1]Division of Molecular Medicine, Boston Children's Hospital, Boston, MA 02115, USA. [2]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA. [3]Codex BioSolutions, Inc., Gaithersburg, MD 20879, USA. [4]The Harvard Cryo-EM Center for Structural Biology, Harvard Medical School, Boston, MA 02115, USA. [5]Department of Biological Chemistry and Molecular Pharmacology, Blavatnik Institute, Harvard Medical School, Boston, MA 02115, USA. [6]Institute for Protein Innovation, Harvard Institutes of Medicine, Boston, MA 02115, USA.
†These authors contributed equally to this work.
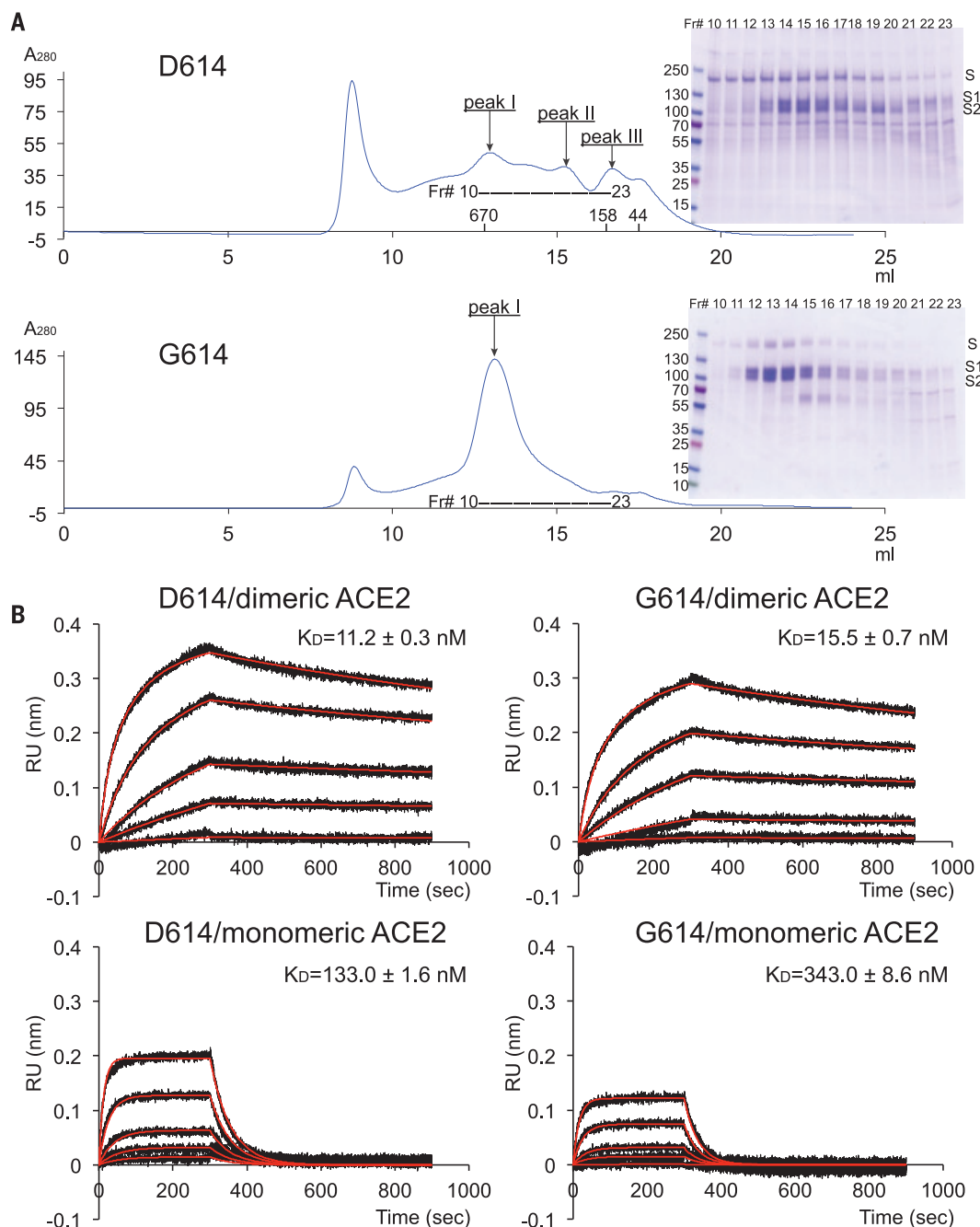*Corresponding author. Email: bchen@crystal.harvard.edu

A

B

Fig. 1. Characterization of the purified full-length SARS-CoV-2 S proteins. (**A**) The full-length SARS-CoV-2 S protein carrying either D614 or G614 was extracted and purified in detergent n-dodecyl-β-D-maltopyranoside (DDM) and further resolved by gel-filtration chromatography on a Superose 6 column. The molecular weight standards include thyoglobulin (670 kDa), γ-globulin (158 kDa), and ovalbumin (44 kDa). Peak I is the prefusion S trimer; peak II is the postfusion S2 trimer, and peak III is the dissociated mono-meric S1. The insets show peak fractions that were analyzed by Coomassie-stained SDS-PAGE. Labeled bands are S, S1, and S2. Fr#, fraction number. (**B**) Binding analysis of fractions of peak I in (A) with soluble ACE2 constructs by BLI. The purified S proteins were immobilized to AR2G bio-sensors and dipped into the wells containing ACE2 at various concentrations (5.56 to 450 nM for monomeric ACE2; 2.78 to 225 nM for dimeric ACE2). Binding kinetics was evaluated using a 1:1 Langmuir binding model for the monomeric ACE2 and a bivalent model for dimeric ACE2. The sensorgrams are in black and the fits in red. Binding constants are also summarized here and in table S1. All experiments were repeated at least twice with essentially identical results. K_D, dissociation constant (binding affinity); RU, response unit.

We determined the cryo–electron microscopy (cryo-EM) structures of the full-length G614 S trimer using RELION (*36*). Three-dimensional (3D) classification identified three distinct classes, each containing a similar number of particles. The three classes represent a closed, three RBD–down conformation; a one RBD–up conformation; and an intermediate confor-mation with one RBD flipped up only halfway. All structures were refined to 3.1- to 3.5-Å resolution (figs. S3 to S8 and table S2). The overall structure of the G614 S protein in the closed, three RBD–down prefusion conforma-tion is very similar to that of our published

D614 S trimer (Fig. 2) (*19*). In the three RBD–down structure, the four domains in each S1, including NTD, RBD, CTD1, and CTD2, wrap around the three-fold axis of the trimer, pro-tecting the prefusion S2. The furin cleavage site is disordered, which makes it uncertain whether this structure represents the uncleaved or cleaved trimer, although the preparation contains pri-marily the cleaved forms (Fig. 1A). The S2 frag-ment folds around a central three-stranded coiled coil that forms the most stable part of the structure; it is also the least variable region among all of the known S trimer structures. The S2 structure is identical in the two G614

structures, with one RBD projecting upward, either completely or partially (fig. S9). In the conformation with one RBD fully up, the two neighboring NTDs, including the one from the same protomer, shift away from the three-fold axis (fig. S9). In the RBD-intermediate con-formation, only the NTD from the adjacent protomer packing directly against the moving RBD shifts. The D614G substitution eliminates a salt bridge between D614 in CTD2 of one subunit and K854 in the FPPR of the adjacent subunit (*19*, *34*), but the FPPR in the three RBD–down conformation of the G614 trimer remains structured.
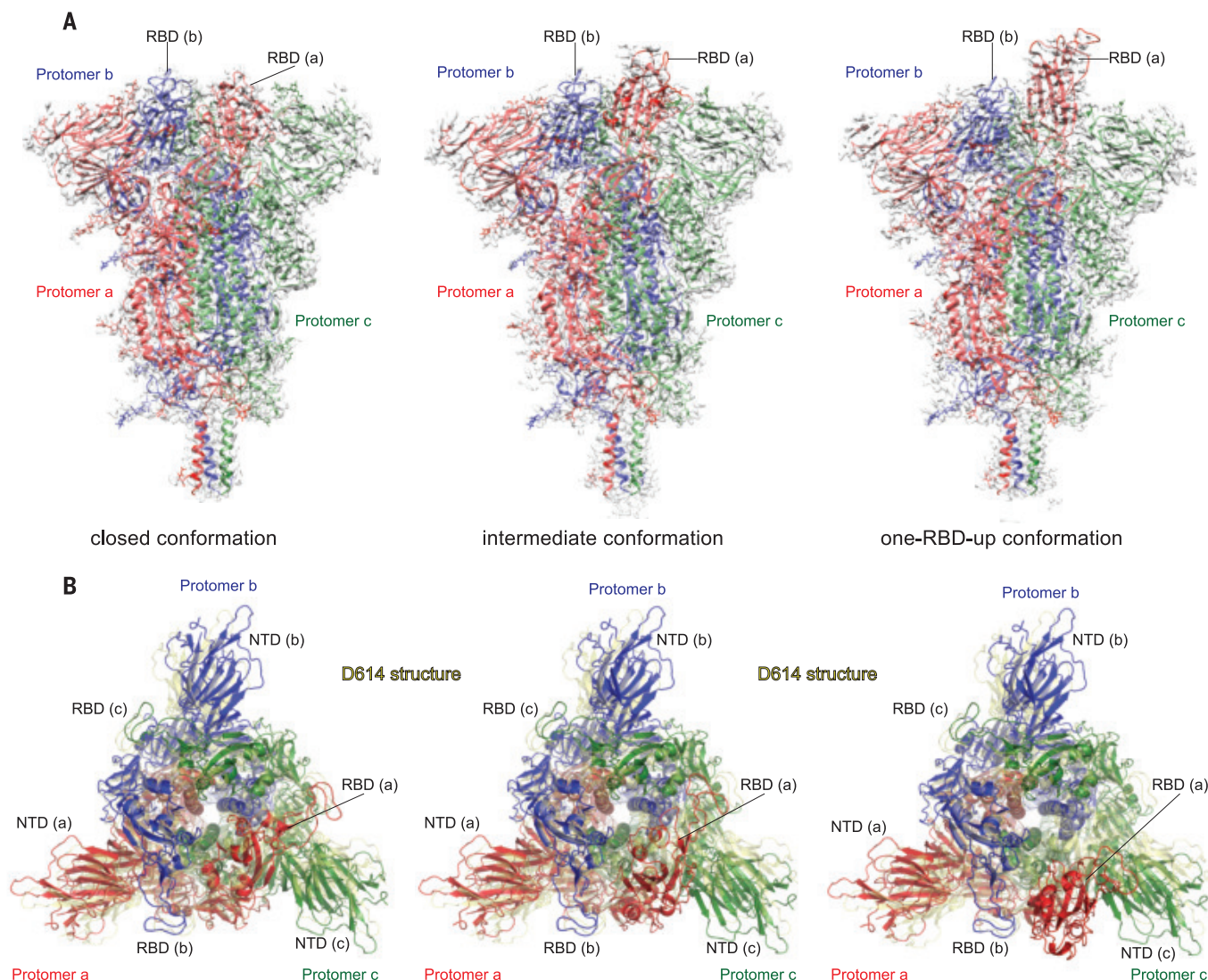
Fig. 2. Cryo-EM structures of the full-length SARS-CoV-2 S protein carrying G614. (**A**) Three structures of the G614 S trimer—representing a closed, three RBD–down conformation; an RBD-intermediate conformation; and a one RBD–up conformation—were modeled on the basis of corresponding cryo-EM density maps at 3.1- to 3.5-Å resolution. Three protomers (a, b, and c) are colored in red, blue, and green, respectively. RBD locations are indicated. (**B**) Top views of the superposition of the three structures of the G614 S in (A) in ribbon representation, with the structure of the prefusion trimer of the D614 S (Protein Data Bank ID: 6XR8) shown in yellow. The NTD and RBD of each protomer are indicated. Side views of the superposition are shown in fig. S8.

To examine the structural changes resulting from the D614G substitution, we superposed the structures of the G614 trimer onto the D614 trimer in the closed conformation, aligning them by the invariant S2 (Fig. 2B). A shift by a clockwise, outward rotation of all three S1 subunits, relative to the D614 structure, is evident even for the G614 trimer in the closed conformation. A similar shift was also observed in the RBD-intermediate and RBD-up G614 structures. Thus, the D614G substitution has led to a slightly more open conformation than that of the D614 trimer, even when all three RBDs are down. The D614G change has apparently also rigidified a neighboring segment of CTD2, residues 620 to 640, which we designate the 630 loop. This loop

inserts into a gap, slightly wider in the G614 than in the D614 trimer, between the NTD and CTD1 of the same protomer (Figs. 3 and 4). The 630 loop is disordered in the closed D614 trimer (fig. S10) because the gap is too narrow for it to insert. The closed D614 trimer thus has three ordered FPPRs and three disordered 630 loops, whereas the closed G614 trimer has three structured 630 loops along with three ordered FPPRs. In the two conformers with one partly or fully open RBD, the two segments are disordered in the RBD-shifted subunit, and their central parts have difficult-to-model densities in one other subunit. The third pair appears well ordered throughout (Fig. 3). Thus, the opening of the RBD in the full-length G614 trimer correlates with a displacement of the

630 loop and the FPPR having moved away from its position in the D614 trimer.

The D614G change did not cause any large local structural rearrangements except for loss of the D614-K854 salt bridge (*19*, *34*) and a small shift of residue 614 toward the three-fold axis (Fig. 4A). The position of the FPPR and the conformation of K854 may allow a hydrogen bond between the K854 amino group and the main-chain carbonyl of G614, perhaps accounting for the subtlety of the structural difference (fig. S11, A and B). Although the loss of the salt bridge involving D614 does not destabilize the packing of the FPPR against the rest of the trimer, it does weaken the FPPR density, especially between residues 842 and 846. The 630 loop, which packs directly against the NTD, CTD1,
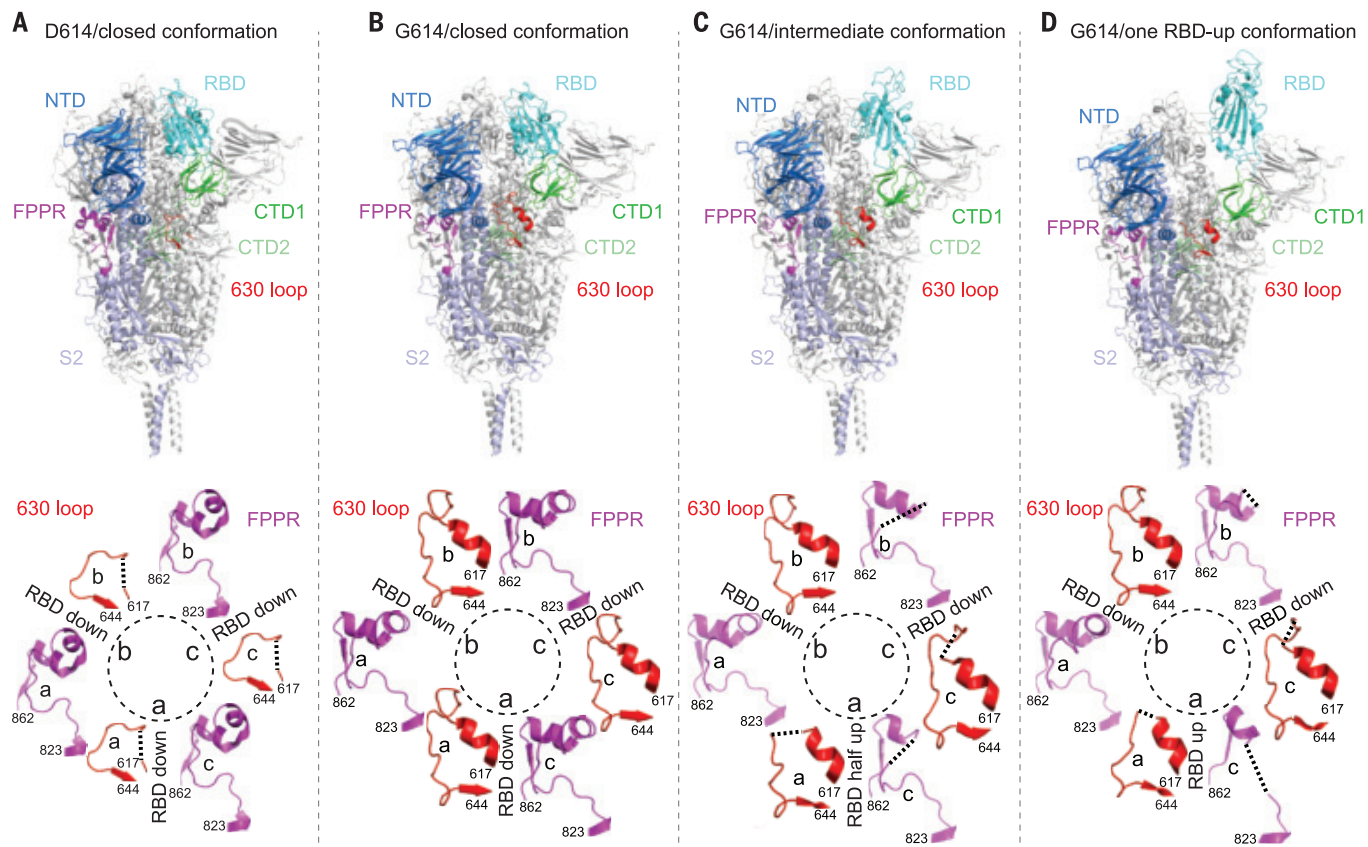
Fig. 3. Cryo-EM structures of the full-length SARS-CoV-2 S protein carrying G614. (**A**) (Top) The structure of the closed, three RBD–down conformation of the D614 S trimer is shown in ribbon diagram with one protomer colored as NTD in blue, RBD in cyan, CTD1 in green, CTD2 in light green, S2 in light blue, the 630 loop in red, and the FPPR in magenta. (Bottom) Structures of three segments (residues 617 to 644) containing the 630 loop in red and three segments (residues 823 to 862) containing the FPPR in magenta from all three protomers (a, b, and c) are shown. The position of each RBD is indicated. (**B** to **D**) Structures of the G614 trimer in the closed, three RBD–down conformation, the RBD-intermediate conformation, and the one RBD–up conformation, respectively, are shown, as in (A). Dashed lines indicate gaps.

and CTD2 of the same protomer, lies close to the S1-S2 boundary of the same protomer and the FPPR of an adjacent protomer (Fig. 4B). Inserting this wedge-like loop between the NTD and CTD1 (Fig. 4C) may help secure the positions of the NTD and CTDs.

CTD2 is formed by two stacked, four-stranded β sheets, with a fifth strand in one sheet contributed by the connector between the NTD and RBD. In the other sheet, an inter-strand loop contains the S1-S2 cleavage site, and thus one strand is the N-terminal segment of S2 (Fig. 4B). In the G614 trimer, one side of the 630 loop packs along a long hydrophobic surface, largely solvent-exposed in the D614 trimer, formed by residues on the upward-facing surface of the CTD2 along with Pro^{295} from the NTD (Fig. 4D). Trp^{633} and Tyr^{636} of the 630 loop appear to contribute to this interaction. S1 dissociation from S2 requires breaking the S2 strand from the second β sheet. An ordered 630 loop that stabilizes the CTD2 by closing off an exposed, hydrophobic surface may retard S1 shedding, thereby enhancing the stability of a cleaved S trimer. We

note that the density for a fatty acid ligand making contacts with the neighboring RBDs in the D614 trimer is absent in all of the G614 reconstructions (fig. S11C) (37), which suggests that the ligand is not required for three RBDs to adopt the down conformation.

To test the impact of the 630 loop on S1 shedding and membrane fusion, we generated five S mutants, each containing a single-residue change either in the 630 loop (W633A, R634E, V635K, and Y636A) or the CTD2 hydrophobic surface (V610K) in the G614 sequence. These mutants expressed the same levels of S, with similar extents of cleavage between S1 and S2, as expected (fig. S12A). When detected by monoclonal antibodies using flow cytometry, mutants V610K and W633A showed markedly lower binding of RBD-specific antibodies [REGN10933 and REGN10987; (38)] and of an NTD-specific antibody [4A8; (39)] than the parental G614 S, whereas binding to an S2-specific antibody [0304-3H3; (39)] was slightly higher (fig. S12B). These results are consistent with the hydrophobic interactions between the 630 loop and CTD2 stabiliz-

ing the cleaved S1-S2 complex and preventing S1 dissociation. The mutant V635K had wild-type phenotypes in these assays, likely because V635 does not make any direct contact with the CTD2. The mutants R634E and Y636A showed intermediate levels of antibody binding because Y636 appears to contribute less to the 630 loop–CTD2 interaction than W633, and R634 may help maintain the loop's overall shape for inserting between domains. Likewise, a similar pattern was observed with these mutants in the cell-cell fusion assay, except that Y636A showed substantially weaker fusion activity than R634E (fig. S12C). Thus, key residues important for stabilizing the S trimer structure appear critical for membrane fusion activity, as premature dissociation of S1 would lead to inactivation of the S trimer.

To further confirm the folding of the 630 loop in the G614 trimer, we collected additional data under slightly different conditions and found the same three classes representing the closed, RBD-intermediate, and one RBD–up conformations (fig. S13A). There is relatively strong density in both 2D class averages
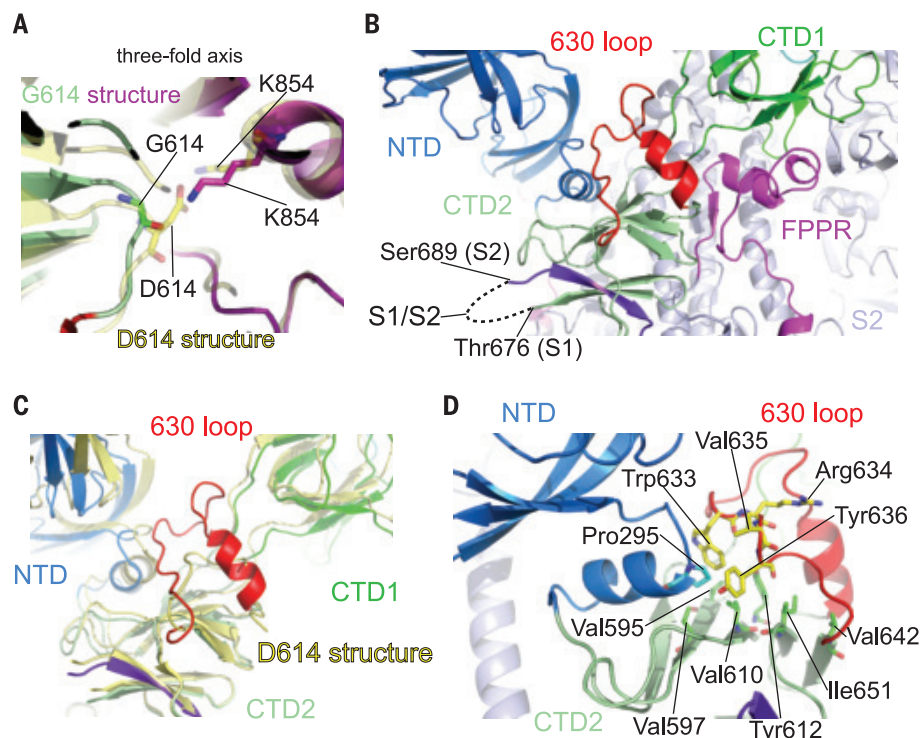
Fig. 4. Close-up views of the D614G substitution. (**A**) A close-up view of the region near the residue 614 with superposition of the G614 trimer structure in green (CTD2) and magenta (FPPR) and the D614 trimer in yellow, both in the closed prefusion conformation. Residues G614, D614, and two K854s from both structures are shown in stick model. The direction of the three-fold axis of the trimer is indicated. (**B**) Location of the 630 loop in the S trimer. The 630 loop is highlighted in red, the NTD in blue, the CTD1 in green, the CTD2 in light green, the S2 in light blue, and the FPPR from a neighboring protomer in magenta. The S1-S2 boundary and the nearest ordered residues Thr[676] from S1 and Ser[689] from S2 are all indicated. A strand from the N-terminal end of S2, packed in the CTD2, is highlighted in purple. (**C**) A view showing that the 630 loop wedges between the NTD and the CTD1 and pushes them apart. (**D**) Packing of the 630 loop against the hydrophobic surface formed by residues Val[595], Val[597], Val[610], Tyr[612], Val[642], and Ile[651] from the CTD2 and Pro[295] from the NTD. Residues Trp[633] and Val[635] from the 630 loop contribute to this interaction.

and 3D reconstructions for the heptad repeat 2 (HR2) region (fig. S1) and detergent micelle, which were invisible in all our previous cryo-EM analyses. The increased length and lack of symmetry limited the resolution of these 3D reconstructions to 4.3 to 4.7 Å. Nevertheless, the density for the 630 loop was evident in the closed trimer even at this resolution (fig. S13B). We note that although S1 in the G614 trimer moves outward from its position in the D614 trimer, the extent of the shift is still appreciably smaller than the shift seen in soluble S trimers stabilized by a trimerization foldon tag and two proline mutations (fig. S14).

Our structures provide an explanation for why the G614 virus, with a more stable S trimer, is more infectious than the original strain (fig. S15). The transition from the closed to the one RBD–up conformation in a G614 trimer requires an order-disorder transition in one 630 loop and a partial disordering of a second. Thus, kinetic barriers will probably make both the forward and reverse transitions slower

than in a D614 trimer, in which all three 630 loops are unstructured in both conformations. In the one RBD–up conformation, S will also shed S1 much more slowly from a G614 trimer than from D614 because the remaining two RBDs are stabilized by the ordered and partially ordered 630 loops, and a return of the first RBD to the down configuration can occur unless locked in place by ACE2 receptor binding. This can account for both the greater prevalence of a one RBD–up conformation and a lower overall ACE2 affinity, because the other two RBDs will remain inaccessible. It can also explain why we captured very few trimers in the RBD-up conformation in our previous cryo-EM study of the D614 trimer but instead saw abundant postfusion S2 (19), because any one RBD–up conformation would proceed to two RBD–up and three RBD–up and shed very quickly, allowing S2 to convert to the postfusion form.

Our interpretation of the structural differences is also consistent with the spike con-

formational distribution on the virions in cryo-ET studies of chemically inactivated SARS-CoV-2. The D614 preparation contains primarily postfusion S2 spikes (24). One study of a G614 virus that had lost the furin cleavage site showed almost no postfusion spikes and a 50:50 distribution of prefusion spikes between fully closed and one RBD–up (21), and another showed 3% postfusion spikes and 97% in the prefusion form (~31%, fully closed; ~55%, one RBD–up; and ~14%, two RBD–up) (22). The structured 630 loop in the G614 trimer not only reinforces the packing among three protomers but also stabilizes the CTD2 to inhibit release of the N-terminal segment of S2, effectively blocking S1 dissociation. This property can account for the paucity of postfusion spikes on the G614 variant.

In addition to the FPPR that might modulate the fusogenic structural rearrangements of S protein (19), CTD2 and the 630 loop within it are probably also the key components of the S fusion machinery. If ACE2 captures the RBD-up conformation (40), expelling both the 630 loop and the FPPR from their positions in the closed S trimer conformation, the FPPR shift may help expose the S2′ site near the fusion peptide for proteolytic cleavage, whereas departure of the 630 loop from the hydrophobic surface of the CTD2 can destabilize this domain and free the N-terminal segment of S2 to dissociate from S1—if the furin site has already been cleaved—and release S1 altogether. Dissociation of S1 would then initiate a cascade of refolding events in the metastable prefusion S2, allowing the fusogenic transition to a stable postfusion structure. This model is similar to that proposed for membrane fusion catalyzed by HIV envelope protein (41).

The SARS-CoV-2 S protein is the centerpiece of the first-generation vaccines that almost all used the D614 sequence. The G614 S trimer is naturally constrained in a prefusion state that presents both the RBD-down and RBD-up conformations with great stability. It is therefore likely to be a superior immunogen for eliciting protective neutralizing antibody responses, which appear to largely target the RBD and NTD (39, 42). It may also be an excellent scaffold for designing next-generation vaccines against new variants that have become resistant to the protections offered by the existing vaccines (43–46).

We suggest that the enhanced infectivity of the G614 virus largely results from the increased stability of the S trimer, rather than the better-exposed RBDs. If the virus that passed from bats to humans, or to an intermediate vector, contained D614 [also present in the bat coronavirus, BatCoV RaTG13 (1)], then it could have gained fitness in the new host by acquiring changes such as G614 for greater stability and infectivity than that observed in

the parental form. Unsurprisingly, the recent fast-spreading variants—including the B.1.1.7 (VUI202012/01; 501Y.V1) lineage from the United Kingdom, the B.1.351 (501Y.V2) lineage from South Africa, and the B.1.1.28 (484K. V2; P.1) lineage from Brazil (*4–6*)—all contain the D614G substitution (table S3), which suggests that the increased transmissibility of the G614 virus has led to a great number of replication events and to greater genetic diversity, despite a lower absolute mutation rate.

REFERENCES AND NOTES

1. P. Zhou *et al.*, *Nature* **579**, 270–273 (2020).
2. F. Robson *et al.*, *Mol. Cell* **79**, 710–727 (2020).
3. B. Korber *et al.*, *Cell* **182**, 812–827.e19 (2020).
4. H. Tegally *et al.*, medRxiv 2020.12.21.20248640 [Preprint]. 22 December 2020. https://doi.org/10.1101/2020.12.21. 20248640.
5. F. Grabowski, G. Preibisch, S. Giziński, M. Kochańczyk, T. Lipniacki, *Viruses* **13**, 392 (2021).
6. C. M. Voloch *et al.*, *J. Virol.* 10.1128/JVI.00119-21 (2021).
7. B. J. Bosch, R. van der Zee, C. A. de Haan, P. J. Rottier, *J. Virol.* **77**, 8801–8811 (2003).
8. M. Hoffmann *et al.*, *Cell* **181**, 271–280.e8 (2020).
9. J. K. Millet, G. R. Whittaker, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15214–15219 (2014).
10. M. A. Tortorici, D. Veesler, *Adv. Virus Res.* **105**, 93–116 (2019).
11. D. Wrapp *et al.*, *Science* **367**, 1260–1263 (2020).
12. A. C. Walls *et al.*, *Cell* **181**, 281–292.e6 (2020).
13. S. M. Gobeil *et al.*, *Cell Rep.* **34**, 108630 (2021).
14. J. Lan *et al.*, *Nature* **581**, 215–220 (2020).
15. R. Yan *et al.*, *Science* **367**, 1444–1448 (2020).
16. J. Shang *et al.*, *Nature* **581**, 221–224 (2020).
17. Q. Wang *et al.*, *Cell* **181**, 894–904.e9 (2020).
18. S. Xia *et al.*, *Cell Res.* **30**, 343–355 (2020).
19. Y. Cai *et al.*, *Science* **369**, 1586–1592 (2020).
20. S. Bangaru *et al.*, *Science* **370**, 1089–1094 (2020).
21. B. Turoňová *et al.*, *Science* **370**, 203–208 (2020).
22. Z. Ke *et al.*, *Nature* **588**, 498–502 (2020).
23. H. Yao *et al.*, *Cell* **183**, 730–738.e13 (2020).
24. C. Liu *et al.*, *Structure* **28**, 1218–1224.e4 (2020).
25. Q. Li *et al.*, *Cell* **182**, 1284–1294.e9 (2020).
26. S. Isabel *et al.*, *Sci. Rep.* **10**, 14031 (2020).
27. J. A. Plante *et al.*, *Nature* **592**, 116–121 (2021).
28. J. Hu *et al.*, bioRxiv 2020.06.20.161323 [Preprint]. 6 July 2020. https://doi.org/10.1101/2020.06.20.161323.
29. D. Weissman *et al.*, *Cell Host Microbe* **29**, 23–31.e4 (2021).
30. E. J. Anderson *et al.*, *N. Engl. J. Med.* **383**, 2427–2438 (2020).
31. L. Zhang *et al.*, *Nat. Commun.* **11**, 6013 (2020).
32. L. Yurkovetskiy *et al.*, *Cell* **183**, 739–751.e8 (2020).
33. T. Zhou *et al.*, *Cell Host Microbe* **28**, 867–879.e5 (2020).
34. X. Xiong *et al.*, *Nat. Struct. Mol. Biol.* **27**, 934–941 (2020).
35. T. Xiao *et al.*, *Nat. Struct. Mol. Biol.* **28**, 202–209 (2021).
36. S. H. Scheres, *J. Struct. Biol.* **180**, 519–530 (2012).
37. C. Toelzer *et al.*, *Science* **370**, 725–730 (2020).
38. J. Hansen *et al.*, *Science* **369**, 1010–1014 (2020).
39. X. Chi *et al.*, *Science* **369**, 650–655 (2020).
40. D. J. Benton *et al.*, *Nature* **588**, 327–330 (2020).
41. M. M. Shaik *et al.*, *Nature* **565**, 318–323 (2019).
42. L. Liu *et al.*, *Nature* **584**, 450–456 (2020).
43. K. Wu *et al.*, bioRxiv 2021.01.25.427948 [Preprint]. 25 January 2021. https://doi.org/10.1101/2021.01.25.427948.
44. C. K. Wibmer *et al.*, *Nat. Med.* 10.1038/s41591-021-01285-x (2021).
45. Z. Wang *et al.*, *Nature* 10.1038/s41586-021-03324-6 (2021).
46. P. Wang *et al.*, bioRxiv 2021.01.25.428137 [Preprint]. 12 February 2021. https://doi.org/10.1101/2021.01.25.428137.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6541/525/suppl/DC1
Materials and Methods
Figs. S1 to S15
Tables S1 to S3
References (*47–56*)
MDAR Reproducibility Checklist

Goverment Of India

INVITATION OF APPLICATIONS FOR THE POST OF DIRECTOR, IIT MADRAS

Ministry of Education

Department of Higher Education

Technical Section - I

Applications are invited for appointment to the post of **Director of Indian Institute ofTechnology (IIT) Madras**. The Director of an IIT is the academic and administrative head of the Institution. He/she is expected to have a minimum of 5 years' administrative experience and leadership qualities to head an Institute of National importance. The candidate/person should be a Ph.D. with first class or equivalent at the preceding degree, preferably in a branch of Engineering. In exceptional cases, candidates with Science, Mathematics or Management degrees may be considered. He/she should have an outstanding academic record throughout and a minimum of 10 years teaching experience as a Professor in a reputed Engineering or Technology Institute or University and should have guided Ph.D. students. The applicant should preferably be less than 60 years of age on the last date of receipt of the applications. The post carries a fixed pay of Rs. 2,25,000/- (Revised) per month, with allowances as per rules.

2. Interested individuals may apply giving their detailed resume in the prescribed format clearly bringing out research, teaching, industry-academia collaborations and administrative achievements, along with a two-page justification in support of their candidature, a two-page vision statement for the institution and contact details of at least two distinguished individuals well acquainted with their work. The application typed in the prescribed format along with enclosures may be sent by Registered/Speed Post to **The Under Secretary (TS.1), Department of Higher Education, Ministry of Education, Room No. 428 "C" Wing, Shastri Bhawan, New Delhi -110 001** so as to reach the Ministry **on or before 31st May, 2021**. The detailed advertisement and the format of application is available on the website (**www.education.gov.in**).

Driving for a difference

I was in the second year of my Ph.D. program when a colleague asked what I would do if I had an extra hour every day. Without much deliberation, I said I would use it to help others. The question haunted me. Like many graduate students, I was overwhelmed with research, teaching, coursework, and some attempt at a personal life. Still, I asked myself, "Do I really need a 25th hour to help other people—or do I need to make better use of the 24 hours I have?" A short time later, I started to volunteer for an organization that provides transportation for senior citizens and people with disabilities. To my surprise, adding this activity to my busy life was just what I needed to calm the chaos I was feeling as a first-generation international graduate student.

I had thought I was ready for the challenges of graduate school. But when I left Nepal to pursue my Ph.D. in the United States, I felt as if I had been taken out of a small pond where I was happy and thriving and dropped into an enormous, enigmatic ocean where every fish swam faster than I could. Even the rules of engagement were unfamiliar. Should I speak out or be quiet? How should I show respect for my professors? How could I connect to my colleagues? Instead of the productive, optimistic, outgoing person I knew myself to be, I was suddenly and uncharacteristically shy. To make matters worse, my research was not progressing at the pace I had anticipated, leading me to put in even more long, mostly fruitless hours—draining my energy with little reward.

I needed something to help me return to my old self. After that conversation with my colleague, I googled "volunteer opportunities near me." A local organization that drives people to stores, appointments, and religious services was looking for volunteers. Having grown up in a bucolic village where everyone knew one another and my grandparents were always close by, I thought serving senior citizens in my new hometown might be just what I needed.

I was a bit anxious before my first ride, as I had not interacted much with people outside school. But my nerves quickly calmed when I picked up my passenger: an older woman on her way to church. She was bemused to see a young international driver and was full of questions about my background. She shared with me how thankful she was to attend worship and get some fresh air. She gave me her blessings, as my own grandparents might have done. Connections like these and the smiles on my riders' faces motivated me to continue driving.

I started to volunteer about 3 hours every weekend—time

"Serving senior citizens in my new hometown might be just what I needed."

I otherwise would have squandered oversleeping or scrolling through social media. Sharing stories with my riders was much more rewarding. What they told me about their lives helped me realize that in every corner of the world, humans are connected with the language of emotions. And seeing how eager my riders were to spend time out and about spurred me to think about how to spend my time, which used to slip away. My previously overwhelming schedule began to feel manageable.

When I shared with one of my riders that I felt mediocre compared with my colleagues who had worked in advanced research facilities, he described how his work abroad had helped him, and he assured me that for my specialty, microbiology, my experience working in a country with a huge burden of infectious disease would be invaluable. I started to reconsider my feelings of inadequacy. As my confidence grew, I began opening up to my fellow graduate students about the ups and downs in my research and was surprised to discover that we were all in the same boat—even my American colleagues who I had perceived as on top of everything. As I developed these connections, my natural personality began to reappear.

I'm now close to completing my degree but still facing some challenges. My adviser moved to a different university and I had to find a new mentor. I am working to publish my research while also applying for jobs. But these hurdles now seem surmountable. I'm proud of who I have become, and I continue to reflect on how I'm using the most valuable thing in life: my time. ∎

Bikash Bogati is a Ph.D. candidate at the University of Tennessee, Knoxville. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER