Combating implicit bias in health care p. 870

Science themes featured at Sundance 2023 p. 874

Organizing pedestrian and other flows p. 923

Science

$15
3 MARCH 2023
science.org
AAAS

TAMING GRAPES

The domestication of an essential vine pp. 880 & 892

EDITORIAL

The first but not the last

After 2 years in Washington, DC, Alondra Nelson is returning to Princeton. A highly decorated sociologist who has written and studied extensively on the intersection of genetics and race, she was appointed by President Joe Biden as deputy director for science and society in the Office of Science and Technology Policy (OSTP) in 2021. The following year, when Eric Lander was removed as the head of that office, Nelson stepped in as its interim director until Arati Prabhakar was named permanent director 8 months later. I recently spoke with Nelson, about a range of issues, from scientific publishing to artificial intelligence. She clearly leaves behind a legacy of science policy-making that encourages equity.

In the world of scientific publishing, Nelson may be best remembered for a memo that set out requirements for providing public access to research findings that have received federal support. The "Nelson memo," as it is commonly referred to, offers different pathways to this end. A scientific paper can be published in an open access journal; alternatively, authors can deposit a paper that has been accepted by a journal (but not yet published) in a public repository. Nelson told me how making "the knowledge ecosystem more equitable in different ways" requires a multifaceted strategy that considers low- and middle-income countries and minority-serving institutions. Options that don't involve article processing charges (as most open access journals do) remove an insurmountable barrier to young researchers and institutions with less funding.

Among her many professional accomplishments, Nelson's academic scholarship and her tenure at OSTP have both served to place science in a social and historical context. "I hope that one of the lessons learned by all of us is that you can't have effective science, or technology, applied science, or innovation, without actually thinking about the use cases," she told me. "The context in which these things will be disseminated and circulated" includes considering "the folks that you need to talk to upstream if you're going to have the downstream outcomes that you want in a particular process, or with a particular product, or scientific output. I would say to scientists…researchers, and engineers who don't have a particular social orientation around their work…that even if they simply care about having efficacious science, it's going to increasingly require them to engage with the social context and the social implications of their work."

> "…the new generation of scientists… understand that the curriculum needs to change…"
> **Alondra Nelson**

Unfortunately, in the United States, the training of scientists has been slow to make room for contemplating the sociology of science. I asked Nelson if she thought there was any hope for getting courses on the social impact of science into undergraduate and graduate education. "I think there is a kind of powerful force, particularly among the new generation of scientists and of researchers…who really understand that the curriculum needs to change, needs to broaden," she said. Nelson noted that the histories of some fields have not been equitable and in some cases have caused harm. Creating policies that "the American public deserves" requires turning this around.

Nelson has been the first Black woman to do so many different things—director of OSTP, head of the Social Science Research Council, and dean of Social Science at Columbia University, among other achievements. I asked how she manages to remain optimistic about the future of science after overcoming the obstacles faced by women and people of color. "The barriers are real…and I would say that science and technology can be more challenging still," she said. "Being the first obviously has its challenges. But I am the first because I go about my work knowing that I'm never going to be the last. Anybody that works with me knows that I go in as the first to just clear space for others…. And so you're the first, but you're not the last…making a change, sometimes too incrementally and too slowly, but it's certainly making a change."

While resuming faculty positions at the Institute for Advanced Study and Princeton University, Nelson will still keep a foot in Washington, DC, at the Center for American Progress to develop ideas about science and technology policy that can be leveraged to improve all of the issues that she has explored throughout her career.

It's likely that we've not heard the last of Nelson being the first.

–H. Holden Thorp

PHOTO: CAMERON DAVIDSON

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

10.1126/science.adh3524

NEWS

> “There is not a consensus in the U.S. government.”
>
> **U.S. National Security Council spokesperson John Kirby**, responding to reports that the Department of Energy now believes “with low confidence” that SARS-CoV-2 emerged from a lab leak in China rather than a natural spillover. Several other agencies favor a natural origin.

IN BRIEF

Edited by **Jeffrey Brainard**

POLICY

U.K.-EU deal opens door to science funds

Researchers in the United Kingdom breathed a cautious sigh of relief this week after the government struck a deal with the European Union to fix post-Brexit disputes including trade across Northern Ireland’s border. The tentative pact, called the Windsor Framework, does not explicitly involve science. But it could end a 2-year delay in finalizing plans for allowing U.K. researchers to apply for grants from Horizon Europe, the European Union’s giant science funding program. In December 2020, the United Kingdom agreed to pay a fee to become “associated” with Horizon Europe, like other non-EU countries including Israel, Norway, and Turkey. But a diplomatic impasse over Northern Ireland—which is part of the United Kingdom but shares a border with EU member Ireland—blocked the arrangement. If the U.K. Parliament approves the Windsor Framework, negotiations for a new deal on Horizon Europe could resume. Even then, some researchers predict they will be long.

Embryo-editing scientist loses visa

SCANDAL | He Jiankui, the Chinese biophysicist imprisoned for 3 years after he edited the genes of human embryos, resulting in three live births, obtained a visa to work in Hong Kong last month—only to see it revoked 10 days later. The 2-year Top Talent visa He received aims to attract those “with rich work experience and good academic qualifications.” In comments on social media and in the local press, He said he hoped to find a position at a Hong Kong university or research institute. Instead, after He’s visa drew attention, Hong Kong officials reconsidered and canceled it, saying he may have made false statements on his application form. They announced they will revise application forms to require disclosure of any criminal convictions. After his April 2022 release from prison, He set up a laboratory in Beijing and has been asking philanthropists to support his research into improving gene therapies for rare diseases. He has not disclosed whether he has found any backers.

A call for geoengineering research

CLIMATE SCIENCE | More than 60 prominent climate scientists this week called for breaking a taboo about solar geoengineering—artificially cooling the planet by making it more reflective—by boosting research on it. Some activists and scientists are staunchly opposed to even studying geoengineering, arguing that it distracts from the necessity of cutting greenhouse gas emissions. But the open letter says decisions on implementing geoengineering schemes are likely in the coming decades, and that simulations and field experiments are needed to understand the schemes’ effectiveness and risks. Among the signatories are retired NASA scientist James Hansen, one of the first to warn about the dangers of global warming, and Harvard University climate scientist David Keith, who has for years tried to gain permission to perform a small-scale geoengineering experiment.

Lab pauses monkey imports

ANIMAL RESEARCH | Charles River Laboratories, one of the largest U.S. importers and suppliers of research monkeys, announced last week it is suspending shipments from Cambodia after receiving a subpoena from the U.S. Department of Justice. In November 2022, the agency indicted members of a smuggling ring that was illegally exporting cynomolgus macaques caught in the wild in Cambodia, labeling them as captive-bred. Charles River said the subpoena is related to several shipments it received from its Cambodian supplier. Charles River said the suspension was voluntary and motivated by “ongoing investigations” of the “supply chain” from Cambodia. The United States is by far the largest importer of the animals globally, mostly for research by pharmaceutical and biotechnology companies. Cynomolgus macaques, which are endangered, accounted for 96% of the nearly 33,000 nonhuman primates the country imported in 2022, according to U.S. government data. About two-thirds of the cynomolgus animals came from Cambodia.

A macaque in China was rescued from smugglers. China has shut down its nonhuman primate exports, increasing demand from suppliers in Cambodia.

PHOTO: CHINA PHOTOS/STRINGER/GETTY IMAGES

ECOLOGY

Whale skin bears record of obscure migration journeys

Scientists used small patches of skin from southern right whales to probe how climate change has shaped their migrations. The technique could help inform conservation measures for the animal, which is recovering from whaling but remains threatened. The species (*Eubalaena australis*) is difficult to track. But the team gathered the skin samples from whales in coastal breeding areas, in part by shooting them with retrievable darts that punch out a small section of skin. The researchers then analyzed chemical signatures—isotopes of carbon and nitrogen—in the skin samples and matched the signatures to isotope ratios mapped across the Southern Ocean over the past 30 years. Whales eat krill and copepods bearing those isotopes, which turn up in fresh whale skin by about 6 months later, creating a record of the whales' past travels. Among the team's findings is that the ocean midlatitudes have consistently remained an important feeding ground. In some parts of the Southern Ocean, the whales are migrating south less often to feed, likely because climate change has reduced populations of krill near Antarctica in some places, the team reports this week in the *Proceedings of the National Academy of Sciences*.

A southern right whale breaches as it plies the Southern Ocean for krill and other food.

PHOTO: FRANCO BANFI/SCIENCE SOURCE

Boron fuel shows fusion promise

ENERGY | Researchers have sparked fusion in a reactor using an alternative fuel mixture that could make potential fusion power plants safer and easier to operate than those burning more conventional fuel. Most experimental fusion reactors use the hydrogen isotopes deuterium and tritium. But tritium is hard to come by, and that fuel combination produces high-energy neutrons that are hazardous to humans and damage reactor walls and components. Alternative fuel made of protons and boron generates no neutrons and produces only harmless helium, but requires a temperature of 3 billion degrees Celsius—200 times the heat of the Sun's core—to burn. Now, a team using a conventional fusion reactor in Japan called the Large Helical Device has reported seeing some fusion reactions at a lower temperature, by using a powerful particle beam to accelerate the protons and help trigger the reactions. The work, reported this week in *Nature Communications*, is far from a practical fusion plant. But a fusion startup, TAE Technologies, which collaborated in the study, hopes to develop one using the fuel.

Cancer's price tag: $25 trillion

BIOMEDICINE | Cancer will cost the world $25 trillion from 2020 to 2050, equivalent to an annual tax of 0.55% on global gross domestic product, a study has found. The analysis estimated treatment costs and the loss of economic productivity by people who become ill or die from 29 types of cancer, accounting for differences across countries in people's education and workforce experience. The most costly cancers include those of the lung, colon, breast, and liver, several of which are also among the most prevalent globally. Increased spending on screening, diagnosis, and treatment could yield substantial health and economic benefits, especially in low- and middle-income countries, which record about 75% of cancer deaths, according to the study, published last week by an international team in *JAMA Oncology*.

NASA science chief named

LEADERSHIP | NASA this week named heliophysicist Nicola Fox as its new science leader. As associate administrator for the agency's science mission directorate, Fox will be responsible for a $7.8 billion budget and more than 100 missions across four divisions: earth science, planetary science, astrophysics, and heliophysics. Fox joined NASA in 2018 to become the heliophysics division chief. Before that, she worked at Johns Hopkins University's Applied Physics Laboratory, where she was project scientist for the $1.5 billion Parker Solar Probe, a mission that is now sampling the Sun's corona in a series of close flybys. Fox replaces Thomas Zurbuchen, who stepped down at the end of 2022.

IN DEPTH

Before deployment, polar researchers—such as these individuals navigating Antarctic icebergs—must pass a thorough medical screening.

PHOTO: KEN KEENAN/NATIONAL SCIENCE FOUNDATION

WORKFORCE

Medical screens for polar research called unfair

History of mental health treatment can be a disqualifier, interviews with *Science* reveal

By **Katie Langin**

A storm is brewing over the U.S. polar research program's medical qualifications process, which screens hundreds of scientists for physical and psychological ailments each year before they deploy to field sites in Antarctica and Greenland where they will have limited access to medical care. Scientists are too often rejected for questionable reasons, some researchers argue. "We have learned of far too many accounts of unfair treatment," a group of senior polar researchers asserts in a letter sent last week to the National Science Foundation (NSF), which funds research in polar regions and coordinates field logistics. "Systems for transparency, reporting, and accountability are needed."

NSF's screening process is administered by the University of Texas Medical Branch (UTMB) and involves detailed medical questionnaires, blood tests, and a physical examination by each applicant's personal doctor. NSF says it is an important tool for catching potentially dangerous health problems before scientists leave for field sites. "The health and safety of all deployed personnel is NSF's top priority," an NSF spokesperson wrote by email. "The medical professionals tasked with clearing deployers are some of the best in the world."

But in interviews with *Science*, several of the 25 letter writers and others expressed concern that some seeking medical clearance—many of them early-career scientists and individuals from underrepresented groups—are failing the screening process for reasons that aren't medically sound. Many of the cases involve scientists who were prescribed medication to treat common mental health conditions, such as anxiety and depression. NSF's medical screening guidelines stipulate that anyone with a psychiatric diagnosis needs to have been "stable" for at least 1 year before their deployment.

But it's not clear what is considered unstable. One graduate student received a notice that she was "not physically qualified" (NPQ) last year after her doctor decreased the dose of her antidepressant medication. Another had the same thing happen after she switched insurance providers and doctors, leaving a short period of time when she didn't have a prescription for her antianxiety medication. (*Science* is not identifying NPQ'd scientists who requested anonymity for fear of future career repercussions.) UTMB did not respond to *Science*'s request for comment about its screening procedures.

It's "really frustrating" that seemingly minor changes in mental health prescriptions appear to be disqualifying researchers from deploying, says Mindy Nicewonger, a climate scientist at Front Range Community College who conducted research in Antarctica as a graduate student and knows one of the students who received an NPQ notice last year. "I also take mental health drugs, and sometimes during seasons you have to up your dose—a lot of people go through seasonal depression," she says.

In another case, a researcher taking antianxiety medication was NPQ'd after moving institutions. A doctor had originally prescribed the medication during the stressful period leading up to her Ph.D. defense. She continued taking it after she moved for her postdoc and started seeing a new doctor. That switch between mental health providers, she believes, is what led UTMB to view her as "unstable." "I later found out that if I had asked anybody in my community, [they] would have told me to lie on the forms about anything mental health

related," she says. "I think that's probably what I would have done in retrospect."

Penalizing researchers for actively managing their mental health is counterproductive, says Seth Campbell, a glaciologist at the University of Maine, Orono. "I've spent over 7 years of my life in the field and [the kind of] person I want in the field with me," he says, is "someone that's willing to actually get mental health support."

Campbell hasn't experienced any problems with the medical screening process himself. But he's witnessed multiple cases where others have run into difficulties, which is why he signed on to the letter sent to NSF. "I want to make sure the students that I'm working with have a better situation—it seems like there's been one issue after another in the past few years."

Some of the scientists who received NPQ notices in recent years ended up going to Antarctica after NSF granted them a waiver. But the NPQs often come just weeks or even days in advance of scientists being deployed, leaving little time for a waiver process many describe as "stressful and confusing." Individuals must sign a document acknowledging that they're "not physically qualified for deployment" and that they absolve NSF and its subcontractors of liability should something go wrong in the field. The individual's employer must sign a similar document, a requirement that often puts scientists in the uncomfortable position of disclosing personal medical information to university officials, including their direct supervisor.

The researchers *Science* spoke with acknowledge the need for a medical screening process. "You definitely should have to be physically qualified in order to go to Antarctica; I don't think anybody's debating that," says University of Kansas, Lawrence, glaciologist Leigh Stearns, lead author of the letter to NSF. But they would like to see more transparency in the entire screening and appeals process and a system for applicants to report allegations of discrimination and abuse to an independent medical ethics board. "If the current system cannot be improved … we petition that another agency or subcontractor be considered for managing polar physical qualifications," the letter authors request.

Michael Gooseff, a hydrologist at the University of Colorado (CU), Boulder, who oversees a project that deploys 31 scientists to Antarctica each year, applauds the letter's recommendations. "A periodic external ethics review is really a great idea," says Gooseff, who wasn't involved in writing the letter. He would also like to see UTMB issue its decisions earlier. "The timing of that kind of stuff prior to deployment is really, really problematic and just overstresses an already stressed system."

Some also worry mental health NPQs could disproportionately harm women—an underrepresented group in polar science—because women are more likely than men to experience symptoms of anxiety and depression. But neither NSF nor UTMB make public any breakdown of the NPQs, which concerns some researchers.

"We're scientists—we want to know," Stearns says. She's been glad to see gains in the number of women researchers in Antarctica since she first started traveling to the frozen continent in 1999. But, "Protections for women in the field have been quite slow to evolve," she adds, referring to pervasive problems with sexual harassment, as recently highlighted in a report released last year. Stearns worries issues with the medical screening process could represent yet another barrier to attracting more women to polar research.

"I suspect if you looked at the stats that you'd see women failing these [medical qualification] exams more than men," says Michael MacFerrin, a glaciologist at CU who received an NPQ notice ahead of a planned deployment to Antarctica in 2020 because he has type 1 diabetes. After he tweeted about his experience last year, other scientists—many of whom were women—reached out to him with NPQ stories of their own. He gave a talk about the issue at the American Geophysical Union's annual meeting in December 2022, which inspired more-senior scientists to write the letter to NSF. They did not ask early-career researchers such as MacFerrin, a research scientist, to sign it in order to shield them from potential career repercussions.

MacFerrin and the letter writers would like to see NSF release aggregate statistics on pass/fail rates by gender, race, and disability status. NSF declined to say whether it was open to releasing such data. "We take the concerns of the community seriously and are working with our prime contractor to evaluate and increase oversight of the physical qualification program," the spokesperson wrote.

The postdoc who received an NPQ notice after changing institutions was able to secure a waiver to travel to Antarctica last year, where she had an "amazing" experience. But she's not sure she'd do it again. "I would have a really hard time convincing myself to try to get to Antarctica again, which is a huge bummer because I study the Antarctic Ice Sheet … and so obviously going there can be a significant benefit for me in my career," she says. "I'm probably better off avoiding the process until they can figure out how to make it right." ■

PHOTO: WIKIMEDIA COMMONS

People across Europe crafted figurines similar to the so-called Venus of Brassempouy.

HUMAN EVOLUTION

Ancient DNA upends European prehistory

Genes reveal striking diversity within similar ice age cultures

By **Andrew Curry**

Thirty thousand years ago, Europe was a land of open steppes with herds of grazing mammoth and other megafauna—and a strikingly uniform human culture. Its inhabitants, whom archaeologists call the Gravettians, dwelled in caves or in shelters built of mammoth bones. They carved palm-size sculptures from mammoth tusk, depicting mammoths, cave lions, and stylized female figurines with elaborate headdresses and exaggerated breasts and buttocks, and left their distinctive art and artifacts from Spain to western Russia. "You can make a case for saying the Gravettian is the first pan-European culture," says University of

related," she says. "I think that's probably what I would have done in retrospect."

Penalizing researchers for actively managing their mental health is counterproductive, says Seth Campbell, a glaciologist at the University of Maine, Orono. "I've spent over 7 years of my life in the field and [the kind of] person I want in the field with me," he says, is "someone that's willing to actually get mental health support."

Campbell hasn't experienced any problems with the medical screening process himself. But he's witnessed multiple cases where others have run into difficulties, which is why he signed on to the letter sent to NSF. "I want to make sure the students that I'm working with have a better situation—it seems like there's been one issue after another in the past few years."

Some of the scientists who received NPQ notices in recent years ended up going to Antarctica after NSF granted them a waiver. But the NPQs often come just weeks or even days in advance of scientists being deployed, leaving little time for a waiver process many describe as "stressful and confusing." Individuals must sign a document acknowledging that they're "not physically qualified for deployment" and that they absolve NSF and its subcontractors of liability should something go wrong in the field. The individual's employer must sign a similar document, a requirement that often puts scientists in the uncomfortable position of disclosing personal medical information to university officials, including their direct supervisor.

The researchers *Science* spoke with acknowledge the need for a medical screening process. "You definitely should have to be physically qualified in order to go to Antarctica; I don't think anybody's debating that," says University of Kansas, Lawrence, glaciologist Leigh Stearns, lead author of the letter to NSF. But they would like to see more transparency in the entire screening and appeals process and a system for applicants to report allegations of discrimination and abuse to an independent medical ethics board. "If the current system cannot be improved … we petition that another agency or subcontractor be considered for managing polar physical qualifications," the letter authors request.

Michael Gooseff, a hydrologist at the University of Colorado (CU), Boulder, who oversees a project that deploys 31 scientists to Antarctica each year, applauds the letter's recommendations. "A periodic external ethics review is really a great idea," says Gooseff, who wasn't involved in writing the letter. He would also like to see UTMB issue its decisions earlier. "The timing of that kind of stuff prior to deployment is really, really problematic and just overstresses an already stressed system."

Some also worry mental health NPQs could disproportionately harm women—an underrepresented group in polar science—because women are more likely than men to experience symptoms of anxiety and depression. But neither NSF nor UTMB make public any breakdown of the NPQs, which concerns some researchers.

"We're scientists—we want to know," Stearns says. She's been glad to see gains in the number of women researchers in Antarctica since she first started traveling to the frozen continent in 1999. But, "Protections for women in the field have been quite slow to evolve," she adds, referring to pervasive problems with sexual harassment, as recently highlighted in a report released last year. Stearns worries issues with the medical screening process could represent yet another barrier to attracting more women to polar research.

"I suspect if you looked at the stats that you'd see women failing these [medical qualification] exams more than men," says Michael MacFerrin, a glaciologist at CU who received an NPQ notice ahead of a planned deployment to Antarctica in 2020 because he has type 1 diabetes. After he tweeted about his experience last year, other scientists—many of whom were women—reached out to him with NPQ stories of their own. He gave a talk about the issue at the American Geophysical Union's annual meeting in December 2022, which inspired more-senior scientists to write the letter to NSF. They did not ask early-career researchers such as MacFerrin, a research scientist, to sign it in order to shield them from potential career repercussions.

MacFerrin and the letter writers would like to see NSF release aggregate statistics on pass/fail rates by gender, race, and disability status. NSF declined to say whether it was open to releasing such data. "We take the concerns of the community seriously and are working with our prime contractor to evaluate and increase oversight of the physical qualification program," the spokesperson wrote.

The postdoc who received an NPQ notice after changing institutions was able to secure a waiver to travel to Antarctica last year, where she had an "amazing" experience. But she's not sure she'd do it again. "I would have a really hard time convincing myself to try to get to Antarctica again, which is a huge bummer because I study the Antarctic Ice Sheet … and so obviously going there can be a significant benefit for me in my career," she says. "I'm probably better off avoiding the process until they can figure out how to make it right." ■

PHOTO: WIKIMEDIA COMMONS

People across Europe crafted figurines similar to the so-called Venus of Brassempouy.

HUMAN EVOLUTION

Ancient DNA upends European prehistory

Genes reveal striking diversity within similar ice age cultures

By **Andrew Curry**

Thirty thousand years ago, Europe was a land of open steppes with herds of grazing mammoth and other megafauna—and a strikingly uniform human culture. Its inhabitants, whom archaeologists call the Gravettians, dwelled in caves or in shelters built of mammoth bones. They carved palm-size sculptures from mammoth tusk, depicting mammoths, cave lions, and stylized female figurines with elaborate headdresses and exaggerated breasts and buttocks, and left their distinctive art and artifacts from Spain to western Russia. "You can make a case for saying the Gravettian is the first pan-European culture," says University of

Tübingen archaeologist Nicholas Conard.

But despite appearances, the Gravettians were not a single people. New DNA evidence, published this week in *Nature*, shows Gravettians in France and Spain were genetically distinct from groups living in what is now the Czech Republic and Italy. "What we thought was one homogenous thing in Europe 30,000 years ago is actually two distinct groups," says Mateja Hajdinjak, a molecular biologist at the Max Planck Institute for Evolutionary Anthropology who was not part of the new study.

The Gravettian data are part of a larger trove of ancient European DNA that reveals striking genetic diversity within apparently unified prehistoric cultures. The sweeping study analyzed 116 newly sequenced genomes and hundreds of previously published ones, ranging from about 45,000 years ago, when the first modern humans reached the continent, to about 6000 B.C.E., and from the Iberian Peninsula to the western steppes of modern-day Russia. It "fill[s] gaps in space and time," says the study's lead author, Cosimo Posth, a geneticist at Tübingen.

The Gravettians, as shown in this reconstruction, had a common culture with sophisticated art and artifacts. But they were two distinct populations.

In period after period, the genetic evidence suggests conclusions drawn from archaeological evidence such as tools, hunting styles, and burial rituals need to be re-evaluated. "These cultural units archaeologists think about as coherent populations don't stand up to the test," says Felix Riede, an archaeologist at Aarhus University who was not part of the study. "It's a major step forward."

Many of the samples were in poor condition and some came from unusual contexts, like the now-submerged landscape between the British Isles and the Netherlands known as Doggerland. New analytical methods and increasingly powerful DNA sequencing tools enabled researchers to squeeze information from extremely degraded bones and teeth, including some that contained just 1% of their original genetic material.

When it comes to the Gravettians, the genetic evidence helps explain subtle regional differences in tool types and subsistence strategies that have puzzled archaeologists for decades. Archaeologists had noted "slight cultural differences, but up till now we didn't know if it was the same or different populations," Hajdinjak says. For example, only people in Eastern and central Europe constructed mammoth bone shelters. University of Leiden archaeologist Alexander Verpoorte, who was not part of the new study, adds, "When you zoom in a little bit, even the female figurines are made in different ways from different materials, deposited in different settings and found in different contexts." Now, it seems they were the handiwork of distinct populations.

The DNA also sheds light on what happened to these ancient Europeans when the climate worsened between 25,000 and 19,000 years ago, a time known as the last glacial maximum when much of Northern and central Europe was blanketed in ice more than 1 kilometer thick. Archaeologists had assumed people including the Gravettians retreated into ice-free areas in southern Europe beginning about 26,000 years ago, then filtered back north several thousand years later as the glaciers melted. That scenario appears to hold true in the Iberian Peninsula and the south of France: People living there before the ice reached its peak persist through the worst of the cold spell, then surge back north and east as the continent warms.

But the Italian Peninsula, long thought to have been a relatively secure refuge, showed something different. Despite what looked to archaeologists like evidence of continuous occupation during and after the glacial maximum, DNA reveals the refuge was actually a dead end. "We expected Italy to be a climate refugium, but there's a sharp and complete turnover—it's a big surprise," Posth says. "The Gravettian population completely disappears." Instead, after the glacial maximum, people in Italy show genetic links to the Near East, suggesting a new population arrived from the Balkans.

About 14,000 years ago, when temperatures across the continent rose sharply in the space of a few centuries, archaeologists recognized cultural changes. But they thought the changes reflected an existing population adapting to hunt in warmer, more heavily forested landscapes. Instead, DNA shows an almost complete population replacement: The people who survived the glacial maximum, known as the Magdalenians, all but vanish and are replaced by populations moving north from postglacial Italy.

The study also looked at the final era of hunter-gatherers in Europe, beginning 10,000 years ago as warming continued to transform the open steppe to dense forests and rich wetlands. Here, again, the genes revealed a surprising wrinkle: Despite broadly similar hunting and gathering lifestyles, people in Western Europe remain genetically distinct from those east of the Baltic Sea.

They even looked different: Genetic data suggest that before the arrival of farmers in northern Europe around 6000 B.C.E., hunter-gatherers in Western Europe had dark skin and light eyes. People in Eastern Europe and Russia, meanwhile, had light skin and dark eyes. Most surprising, despite the lack of geographic barriers between modern-day Germany and Russia, the two groups spent millennia not mingling. "From 14,000 years ago to 8000 years ago, they do not mix at all," Posth says. But he acknowledges that the team's samples don't cover the continent completely, and the likely contact zones—in Poland and Belarus, for example—lack samples. More genetic data from those areas might show the two populations mixing locally.

Archaeologists are expected to welcome the new genetic data, even though they may force many to re-examine old ideas, says Jennifer French, an archaeologist at the University of Liverpool who was not part of the study. "This genetic data shows we've oversimplified what was going on in terms of population interaction," she says. "It provides a lot more nuance than we've been able to with archaeological data alone." ■

ILLUSTRATION: TOM BJOERKLUND

Compared with the normal leaf (left) of a tobacco relative, a leaf engineered to make an antibodylike receptor defends itself against an experimental virus.

PHOTOS: ANDRÉS POSBEYKIAN/SAINSBURY LABORATORY

AGRICULTURE

Antibody-based defense may protect plants from disease

A "creative" strategy to keep crops healthy borrows key pathogen detectors from the animal immune system

By **Erik Stokstad**

COVID-19 has tragically given many people a crash course in the importance of antibodies, pathogen-targeting proteins produced by the sophisticated immune systems of humans and other animals. Now, researchers from a U.K. plant research institute have found a way to endow plants with an antibody-based defense for a specific threat, potentially speeding the creation of crops resistant to any kind of emerging virus, bacterium, or fungus.

"It's a really creative and bold approach," says Jeff Dangl, a plant immunologist at the University of North Carolina, Chapel Hill. Roger Innes, a plant geneticist at Indiana University, Bloomington, adds: "This would be much, much faster than standard plant breeding and hopefully much more effective."

The strategy is to inoculate an alpaca or other camel relative with a protein from the plant pathogen to be targeted, purify the unusually small antibodies they produce, and engineer the corresponding gene segment for them into a plant's own immune gene. In a proof of concept described on p. 934, this approach equipped a model plant species with immunity against an engineered version of a virus that infects potatoes and related crops.

Farmers lose many billions of dollars to plant diseases each year, and emerging pathogens pose new threats to food security in the developing world. Plants have evolved their own multipronged immune system, kick-started by cell receptors that recognize general pathogen features, such as a bacterial cell wall, as well as intracellular receptors for molecules secreted by specific pathogens. If a plant cell detects these molecules, it may trigger its own death to save the rest of the plant. But plant pathogens often evolve and evade those receptors.

A long-standing dream in plant biotechnology is to create designer disease resistance genes that could be produced as fast as pathogens emerge. One approach is to edit the gene for a plant immune receptor, altering the protein's shape to recognize a particular pathogenic molecule. This requires specific knowledge of both the receptor and its target on the pathogen.

Instead, Sophien Kamoun, a molecular biologist at the Sainsbury Laboratory, and his colleagues harnessed an animal immune system to help make the receptor modifications. During an infection with a new pathogen, animals produce billions of subtly different antibodies, ultimately selecting and mass-producing those that best target the invader.

Camelids, which include alpacas, camels, and llamas, are workhorses for antibody design because their immune systems create compact versions, called nanobodies, encoded by small genes. As a proof of principle of the new plant defense strategy, Kamoun's group turned to two standard camelid nanobodies that recognize not pathogen proteins, but two different fluorescent molecules, including one called green fluorescent protein (GFP). The team chose these nanobodies to detect test viruses, in this case a potato virus, engineered to make the fluorescent proteins.

Jiorgos Kourelis, a postdoc in Kamoun's lab, first melded the gene for the GFP-targeting nanobody to the gene for an intracellular immune receptor in the tobacco relative *Nicotiana benthamiana*. In a follow-up demonstration, he repeated the feat with the gene for the nanobody recognizing the other glowing protein. It took several tries and tweaks to create plants that did not mount autoimmune responses because of the modified receptors, which would have stunted growth and impaired fertility.

Next, Clémence Marchal, also a postdoc in Kamoun's lab, investigated how well plants with the nanobody-enhanced receptors detected the altered potato viruses. Marchal found that the plants mounted a vigorous immune response—the patches of self-destructing cells were visible to the naked eye—and experienced almost no viral replication, whereas leaves from control plants suffered from infection.

Plant breeders often "stack" resistance genes into plant varieties to add protection against several diseases at once. In the team's experiment, plants given genes for both kinds of nanobodies were protected against either viruses. "The exciting part about this technology is we have the potential of made-to-order resistance genes and keeping up with a pathogen," Kamoun says.

The group has since engineered a crop to produce nanobodies that detect actual pathogen molecules, although Kamoun declines to identify the plant before the team has tested whether it withstands assault by the pathogens. The Sainsbury Laboratory has filed patent applications worldwide on the strategy, including in Europe, where public opposition to genetic engineering means it is unlikely to be commercialized any time soon. But Kamoun says there is commercial interest from elsewhere.

Dangl and others are optimistic that the nanobody approach should work in crops. "This technology is a potential game changer," he says. Ksenia Krasileva, a geneticist at the University of California, Berkeley, says the fusion of nanobodies with plant immune receptors opens up a vast body of biomedical knowledge for plant scientists. "We can now tap into all of that research and translate it to save crops. We have a perfect merging point here." ■

White-tailed deer are so numerous that they outrank all other wild land mammals in biomass.

PHOTO: JASPER DOEST/MINDEN

CONSERVATION

Wild mammals add up to a 'shockingly tiny' total biomass

Humans and domestic species far outweigh other mammals

By **Elizabeth Pennisi**

What wild mammal treads most heavily on the land? Not elephants, according to a new global estimate of the total masses of mammal species. Not wild mice, despite their numbers. The heavyweight champion is that furtive denizen of parks, meadows, and forests throughout the Americas, the white-tailed deer. It accounts for almost 10% of the total biomass of wild land mammals.

The study, which honed its numbers using artificial intelligence, "is the first that provides quite convincing values for mammals," says Patrick Schultheiss, a behavioral ecologist at the University of Würzburg. Published this week in the *Proceedings of the National Academy of Sciences*, it concludes that wild land mammals alive now have a total biomass of 22 million tons, and marine mammals account for another 40 million tons.

Those numbers are relatively puny: Ants alone amount to 80 million tons, Schultheiss has estimated. But the comparison the team hopes will capture attention is with humans, who weigh in at 390 million tons, with their livestock and other hangers-on such as urban rats adding another 630 million tons. It is stark evidence of how the natural world is being overrun, researchers say. "I hope it will be a wake-up call to humanity that we should do all we can do to conserve wild mammals," says lead author Ron Milo, a quantitative biologist at the Weizmann Institute of Science.

Milo says he and his colleagues have long believed numbers can "provide a sixth sense of the world." In 2018, they grabbed headlines by estimating the global weight of all life; 2 years later, they added the global weight of all humanmade objects and infrastructure, from cars to buildings. They also made a rough estimate of 50 million tons for wild mammals—"a shockingly tiny fraction of the mass of life on Earth," recalls Shahid Naeem, a biodiversity ecologist at Columbia University. Since then, the team has worked to sharpen its estimate of this fraction.

At Milo's lab, Lior Greenspoon and Eyal Krieger were able to find detailed data on the global number, body weight, range, and other measures for 392 wild mammal species, enough to calculate their total biomass directly. To predict the total mass for less-studied mammals, they used the data for half of the 392 species to train a machine learning system. They tested and refined the model until it could accurately predict the biomass of the other half of the species. They then fed whatever data they could find—ranges, body sizes, abundance, diets—for each of about 4400 additional mammal species into the model to estimate their biomasses and abundance. The work "was a major undertaking," says Renata Ivanek, a veterinarian and epidemiologist at Cornell University who has evaluated ways to estimate the biomass of livestock.

On land, much of the wild mammalian biomass is concentrated in a few large-bodied species, including boar, elephants, kangaroos, and several kinds of deer. The top 10 species account for 8.8 million tons—40% of the estimated global wild land mammal biomass, Milo's team reported. Rodents—not counting human-associated rats and mice—make up 16% and carnivores account for 3% of that biomass. Among marine mammals, baleen whales account for more than half of the biomass. But in sheer numbers, bats rule the mammalian world: They may constitute two-thirds of individual wild mammals, though only 7% of the total terrestrial mass.

In contrast, on the domesticated front, cows collectively weigh 420 million tons and dogs about as much as all wild land mammals, the new study reports. The biomass of housecats is about double that of African elephants and four times that of moose.

"There is much uncertainty around the estimated [biomasses]," Ivanek adds, "but it is a start." The contrast with the masses of wild animals is having Milo's desired effect. These results "changed my notion about the seemingly endless ubiquity of wildlife and provided insight into the extent to which our activity as humans has impacted our world," says Ece Bulut, a Cornell food scientist who collaborates with Ivanek.

But so what? It is an "impressive effort to provide a snapshot of the state of the mammalian world," Naeem says, yet he contends that this snapshot "won't drive conservation nor transform our way of thinking about the issues."

Not true, counters Sabine Nooten, an insect ecologist and Schultheiss's collaborator at Würzburg. "We can only conserve what we understand, and we can only truly understand what we can quantify." ■

FUNDING

Wealth gap in biomedical research is growing

Women and Black researchers are less likely to hold multiple NIH grants

By **Jocelyn Kaiser**

Researchers who hold at least three grants from the National Institutes of Health (NIH) make up a growing portion of its grantees, according to a new study. White men predominate among these "super" principal investigators (PIs), with women and Black researchers much less likely to be part of this elite group.

The trends are "concerning" both because of the concentration of resources in a relatively small number of labs and because of the inequities they reflect, say the authors of a study published this week in *JAMA Network Open*. Others agree. "A more level playing field where there were fewer superwealthy PIs would allow more excellent scientists to stay funded," says University of North Carolina, Chapel Hill, cell biologist Mark Peifer, who has supported capping the amount of NIH support allotted to individual investigators. "But if we have a system where there is an elite, let's make that elite represent the broader community."

The new study, led by physician and health equity researcher Dowin Boatright of New York University's Grossman School of Medicine, analyzed grant data for the nearly 34,000 investigators NIH funded in 2020, homing in on the 11.3%—nearly 4000—who held three or more grants, totaling a median of $1.3 million per grantee. The relative size of this group has tripled since 1991, when only about 3.7% of the then-18,820 investigators had three grants or more. Even when NIH's budget expanded in the early 2000s and since 2015, enabling it to fund more grants, the elite fraction rose, notes Yale University M.D.-Ph.D. immunobiology student Mytien Nguyen, the paper's first author. "Whenever there's an uptick in NIH funding, that creates more inequity" in the form of a larger share of super-PIs, she says.

Also troubling are the demographics of these super-PIs. White men, who made up 64.7% of all investigators, were overrepresented, accounting for 73.4% of the super-PIs. Black researchers made up 1.8% of all investigators but just 0.9% of elite PIs. Black women were one-third as likely to be an elite PI as white men; just 12 Black women held at least three grants. (Asian men, by contrast, made up 15% of all funded investigators but 18.7% of the elite group.)

NIH's extramural grants chief, Michael Lauer, notes the agency has published similar data in the past couple of years. "Older white men are more likely to have more grants, and that differential if anything seems to be widening a bit," Lauer says.

NIH has noted that studies suggest productivity per grant dollar levels off after two grants. But when the agency proposed limiting investigators to the equivalent of three basic R01 research grants in 2017 to free up funds for more investigators, the proposal was shot down by critics who argued it would cripple large, productive teams.

Lauer points to one encouraging trend: The share of NIH's total grant dollars going to the top 1% of grantees has declined slightly, from 11% a few years ago to about 10% in 2020. He thinks this may be related to a 2017 NIH policy giving preference to applications from early-stage investigators, which has led to steady rise in the number of grants awarded each year to this group.

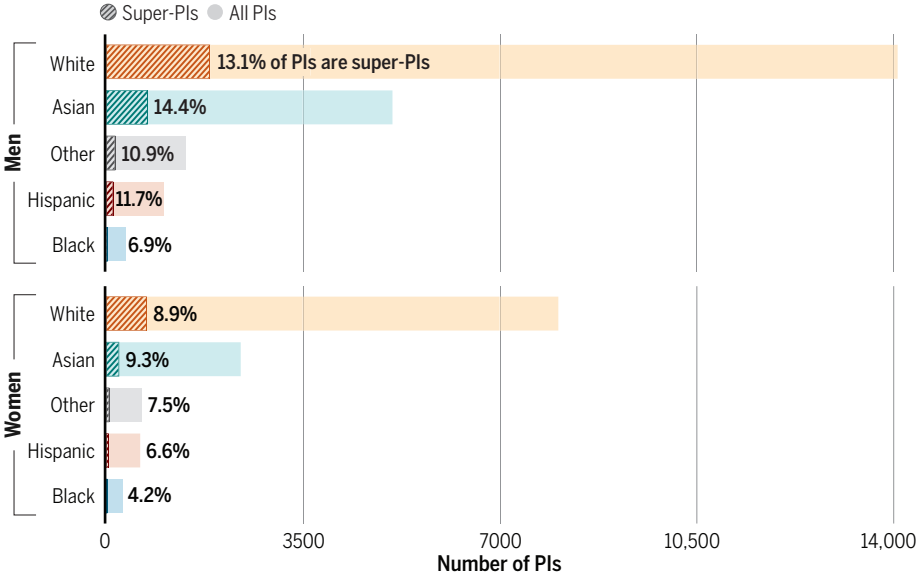
But there's still a long way to go. Both the study team and NIH point in part to the struggles of women and racial and ethnic minorities to climb the academic ladder. Those at the top "may have access to more resources, which enabled them to apply successfully for more grants," Lauer says. NIH should expand mentoring programs aimed at helping women and minorities write more competitive proposals, the new paper's authors suggest.

Nguyen and Boatright also say NIH should diversify the makeup of the scientists who sit on grant peer-review panels—a step NIH is already taking, Lauer says. And they think a plan to enhance diversity, for instance by training minority students, should be a scored component of grant proposals. NIH is pilot testing this idea as part of its BRAIN neuroscience initiative.

NIH's efforts to help women and minorities need to be "multifaceted," says neuroscientist and Associate Vice Provost Sherilynn Black of Duke University, who leads initiatives for faculty advancement at Duke. "This study demonstrates that we are not where we need to be if we truly want science to be as equitable and innovative as possible." ■

Unequal share

Overall, 11% of all principal investigators (PIs) with National Institutes of Health funding in 2020 held three or more grants. However, the numbers of these "super-PIs" vary by gender and race and ethnicity. White and Asian men were disproportionately more likely to belong to this elite group, and Black women were least likely.



CREDITS: (GRAPHIC) K. FRANKLIN/*SCIENCE*; (DATA) M. NGUYEN *ET AL.*, *JAMA NETW. OPEN*, 6(2):E230855 (2023)

FEATURES

DO NO UNCONSCIOUS HARM

Implicit bias can degrade health care. Now, researchers are finding new ways to recognize and mitigate it

By **Rodrigo Pérez Ortega**; *Illustrations by* **Thumy Phan**

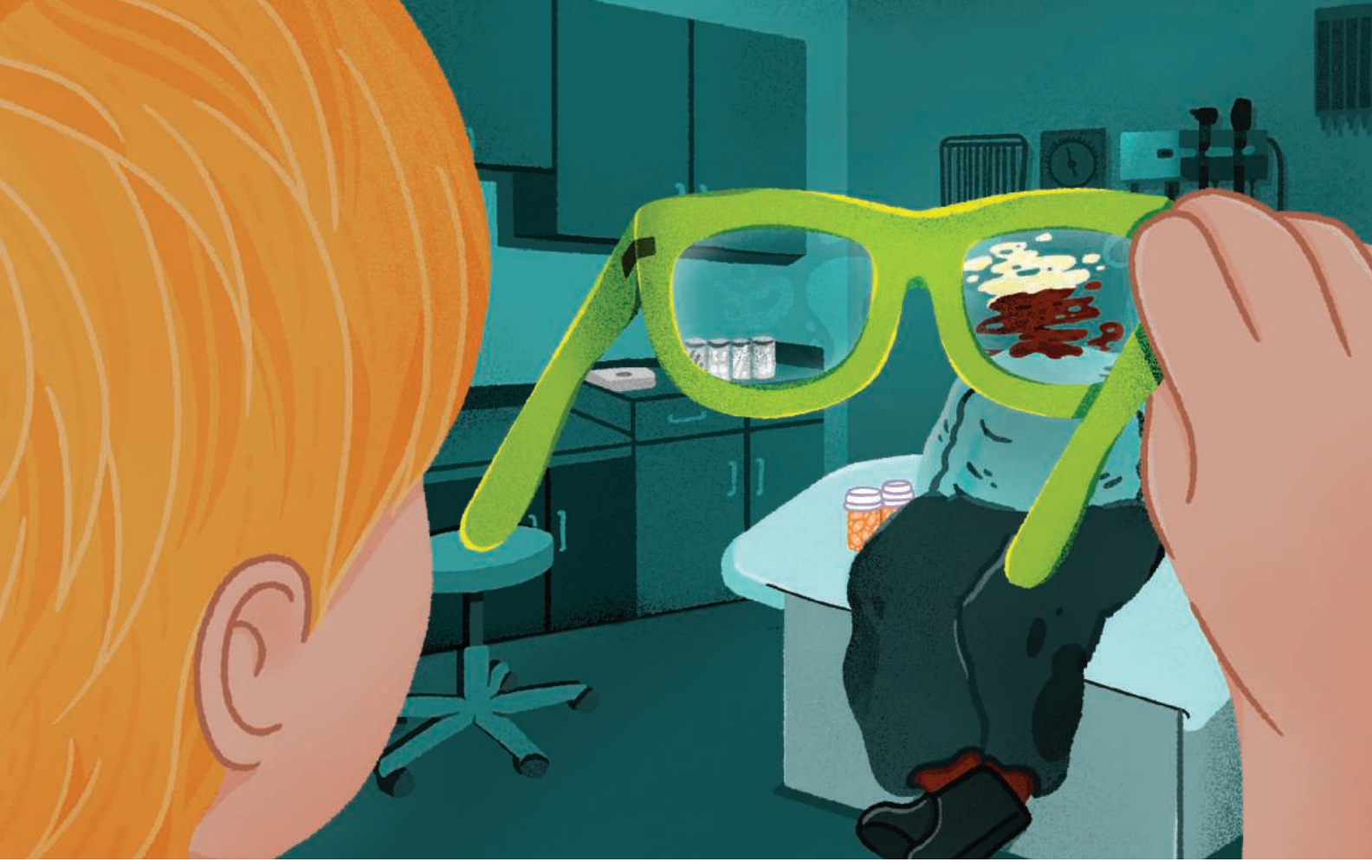
Before being diagnosed with an autoimmune disease several years ago, Linda Chastine says she was skinny. But the lifesaving steroid treatment for her condition made the Seattle community organizer gain roughly 45 kilograms in 1 year. Yet her doctors often fail to recognize the relationship between the autoimmune condition, her medication, and the weight gain. Instead they routinely bring up being overweight as her main problem, she says, and suggest she count calories. "It's a very traumatic experience to have a doctor dismiss a lot of your complaints or concerns based on weight," Chastine says.

Stories like Chastine's are unfortunately common, say researchers who examine how implicit biases—unconscious assumptions based on skin color, gender, sexual preference, or appearance—in health care providers affect patient care. Chastine, who is Black and queer, is now channeling her troubled experience with the medical establishment to aid studies of implicit bias and identify ways to counter it. She is part of a 5-year collaboration between various departments at both the University of Washington (UW) and the University of California, San Diego (UCSD), in which a team is developing a tool to give physicians feedback in real time during patient visits—or shortly after—on what they can do to mitigate their unconscious prejudices.

That project, called UnBIASED (Understanding Biased patient-provider Interaction and Supporting Enhanced Discourse), is at the leading edge of a wave of efforts to counter the negative effects of bias in medicine. From creating new models of education and training, to developing accurate tests to objectively measure pain, scientists are working to provide health care workers—and institutions—with the tools to diminish bias and provide equitable care. Although it might be too early to know whether these interventions are successful and long-lasting, some strategies appear promising.

"We're just looking at one little slice" of how to eradicate implicit bias in health care, says Andrea Hartzler, the UW biomedical informatician who leads UnBIASED, which is funded by the National Library of Medicine. "It's going to take a tool chest of all kinds of different interventions." One key component, she adds, is to go beyond the individual doctor or nurse and target structural issues within their institutions that promote these biases.

EVERYONE HAS prejudices that affect how they perceive and behave with others. And although many people might be aware of some—their ex-



plicit or conscious biases—and intentionally try to compensate for them, other hidden ones still lurk and can influence attitudes and interactions.

These implicit biases are widespread among health care providers, as Janice Sabin discovered in the late 2000s. In her research back then—as a social welfare Ph.D. student at UW—Sabin had asked 95 doctors from the Department of Pediatrics at Seattle Children's Hospital to take a test that would determine whether they had a "hidden" bias toward a certain race. "I was terrified," Sabin, now a biomedical informatics professor at UW, recalls. "This wasn't just asking them questions about bias and racism, this was actually going into their mind."

Sabin used the well-known Implicit Association Test (IAT), which determines how strongly an individual associates a trait—such as race or sexual orientation—with a subjective value, such as "good" or "bad." The quicker you match each concept to a subjective value, the greater the association and the higher your score, which broadly indicates a stronger implicit association between the trait and value.

Sabin found the doctors she tested—a few of them nonwhite—had a clear, unconscious preference for white people over Black people. It was one of the first studies showing health care providers had unintentional racial prejudices. "It was kind of scary because this was a concept completely foreign to [many] people at the time," Sabin says.

The IAT remains a standard tool for measuring implicit bias, although some have criticized it because it has to be taken several times to reveal a reliable result, as people's scores could change every time they take it. Even when people come out neutral on race, most studies will reveal some kind of unconscious prejudice, such as an unrecognized preference of certain sexual orientations or religions. "We all have some kind of hidden bias," Sabin says. Not only are these biases present among health care providers, but research suggests they are likely affecting diagnoses and treatment decisions, and in turn contributing to health disparities affecting people of color, women, and members of the LGBTQ+ community and other historically marginalized groups.

A 2020 study by Rachel Hardeman, a reproductive health equity researcher at the University of Minnesota's Center for Antiracism Research for Health Equity, and colleagues showed Black newborns are twice as likely to die in the care of a white physician than a Black doctor, for instance. Another study from 2022 found women and people of color with chest pain wait longer to be treated in the emergency room compared with white men.

Pain assessment by health care providers has become a fertile ground for research into unconscious medical prejudices and a classic example of the way bias undermines the care of minorities. Implicit bias often shows up when there's no objective test or measurement for a symptom, and that's the case for pain. "Pain is subjective and how people overtly show signs of pain varies across different cultures, across gender identities," says Kristyn Smith, an emergency physician at the University of Pennsylvania.

Previous studies have shown physicians tend to underestimate pain experienced by women and people of color and discount their complaints. In mock medical cases, reported in 2016 in the *Proceedings of the National Academy of Sciences*, white medical residents estimated Black patients felt less pain than white ones, and as a result, made less effective treatment recommendations for Black people. But this bias isn't limited to doctors: A study from 2021 showed lay people underestimate the pain in women patients compared with men pa-

ILLUSTRATION: THUMY PHAN

tients, opting to treat women with psychotherapy and men with painkillers.

"My personal experience is that it's not something that is due to people necessarily having [overt] prejudices or wanting to do a bad job," says Indiana University, Indianapolis, psychiatrist Alexander Niculescu III. "It's just that [when] lacking objective tools, sometimes people can make assumptions that are [going] in the wrong direction."

Implicit biases can also affect health outcomes simply because patients feel discriminated against and don't come back to their doctor. A 2017 systematic review of research on implicit bias in health care providers examined 42 published studies conducted mainly in the United States but also in nine other countries, and concluded there was "a significant positive relationship between level of implicit bias and lower quality of care."

Many other inequities, such as poverty and redlining, prevent certain groups from accessing good health care, Smith says. "If you include the biases of your health care professionals, then that creates a perfect storm for health care disparities to continue."

SCIENTISTS HAVE long studied several kinds of interventions that attempt to "erase" implicit bias, but few of them have shown lasting effects. "There is a robust science around implicit bias," Hardeman says. But, "There is no gold standard for how to intervene right now. It's imprinted in our brains in ways that make it really hard."

Simple interventions can dampen biases, as measured by successive IATs, but the changes are usually modest and don't persist. In a 2001 experiment, for example, researchers showed images of admired Black people—such as Denzel Washington or Colin Powell—versus disliked white individuals—such as Jeffrey Dahmer and Howard Stern—to study participants, and saw that this exposure significantly weakened a pro-white preference for 24 hours, but not for much longer.

Simply asking health care providers to take the IAT without providing context or tools can be counterproductive. A study in 2015 indicated that when medical students are told about their unconscious bias without direction on overcoming it, they tend to get anxious, confused, and nervous interacting with patients who belong to social groups different from their own. That's why even a quick training on skills to mitigate implicit bias can go a long way, according to Hardeman.

So, Hartzler and others are developing feedback tools that will help clinicians confront and make sense of their biases. The team conducted interviews with a small group of primary care doctors to get a sense of the best way to provide feedback to providers on implicit bias. The researchers also spoke with traditionally marginalized groups, including people of color and LGBTQ+ individuals, to learn about biased behaviors that may be on display in patient-doctor interactions. The UnBIASED team then recruited physicians and "community champions," including Chastine, to help design culturally sensitive experiments that could reveal personal biases. "It's really great to have all these minds that want to be innovative about how we are addressing bias," Chastine says.

Brian Wood, an infectious disease clinician with UW and Harborview Medical Center, is one such physician volunteer with UnBIASED. Wood, who primarily sees people with an HIV diagnosis, says his physicians group serves a diverse population who often feel stigmatized by doctors. "I often hear from Black patients how they feel discriminated against," says Wood, who is white. "And my transgender patients often express how fearful they are of seeking care from any provider they don't know," because of painful past encounters with the medical establishment, he says.

> **"There is a robust science around implicit bias. [But] there is no gold standard for how to intervene right now."**
>
> **Rachel Hardeman**, University of Minnesota's Center for Antiracism Research for Health Equity

That made him eager to take part in UnBIASED's first experiments, which rely on cameras installed in exam rooms. The cameras in Wood's Seattle clinic captured interactions between him and his patients, including close-ups of his and their facial features and body language. "I found quite quickly that the patient and I both forgot the cameras were there and just fell into our usual routine and conversation," he says.

The UnBIASED team then used a type of artificial intelligence (AI) known as machine learning to analyze patterns in the recordings and identify nonverbal cues that could indicate implicit bias. In one of the clips Wood was later shown, he was talking with a patient while leaning forward with his arms crossed on the desk, body language he worries may have made him seem closed and unapproachable. "I reflected on my own as to how that body language might be felt and perceived by the patient," he says. Wood, who hopes to improve his demeanor, says he welcomed such feedback and is eager for more.

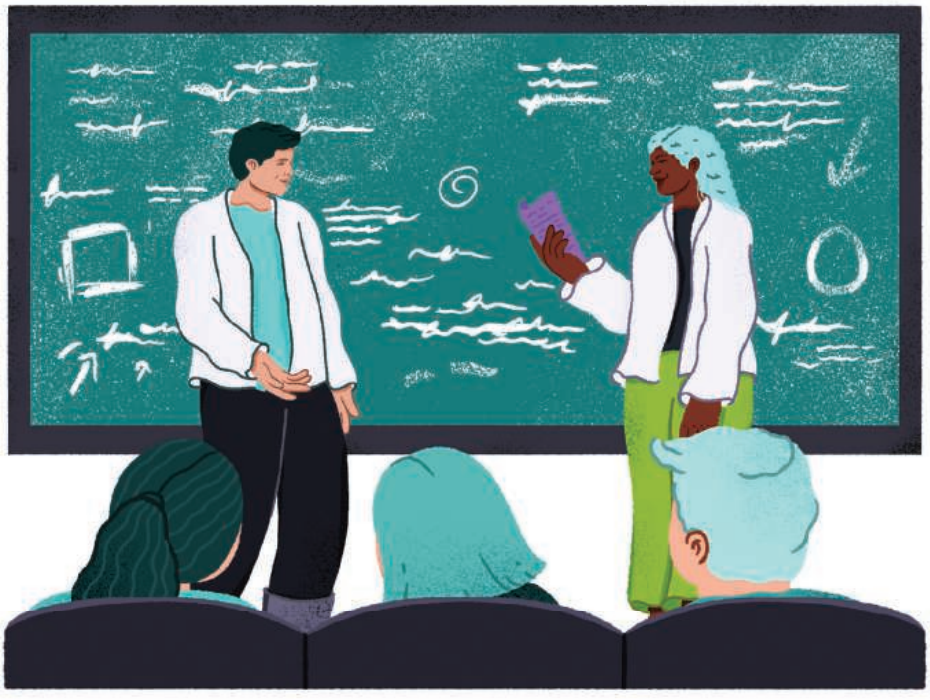
"Reflecting on possible negative moments during a visit was not easy, but felt important and valuable," Wood says.

The team is now working on translating the experiment's results into feedback like "digital nudges"—such as an icon that pops onto the computer screen, a wearable device, or other mechanism telling physicians to interrupt patients less or look them in the eye more often. But the UnBIASED team still has challenges interpreting the data in the recordings. For instance, nonverbal signals are nuanced, Hartzler says. "It's not always as simple as 'more interruptions means bad.'"

Others using computer software to research implicit bias in medicine are also struggling to give physicians meaningful feedback. Nao Hagiwara, a social and health psychologist at Virginia Commonwealth University, and her team are analyzing dozens of nonverbal and verbal communication behaviors, such as facial expressions and voice changes, in recordings of primary care physicians' interactions with people who have type 2 diabetes. Their software hasn't yet identified behaviors that could clearly be linked to bias or had an adverse effect in the patient's outcome. One reason for this murkiness, Hagiwara suggests, is that multiple different cues likely interact to influence patient outcomes whereas studies so far tend to analyze the impact of only one behavior at a time.

Smith is working on a different type of implicit bias intervention: creating clinical simulations from her time working in a trauma center that served a majority Black population. Such emergency departments, where doctors and nurses are often overworked and in a high-stress environment, are ripe for implicit bias to kick in easily, Smith says. She recalls an instance when a Black man arrived at the hospital with a gunshot wound. He died, and shortly after, a social worker came up to her and said, "He was always here … he was just shot 2 weeks ago. … I wonder what he did this time."

Smith recalls that the social worker was normally a fierce advocate for patients, but says that comment was steeped in bias



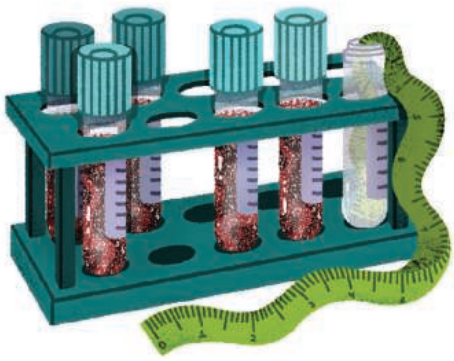
because it suggested the gunshot victim had done something wrong and deserved his fate—an attitude that could affect care. But a gunshot victims' actions should not affect their care, Smith stresses. "Patients who get shot are victims first, and deserve to be treated with respect and sensitivity." (The social worker was very remorseful after Smith told her that her comment was inappropriate.)

Using this experience, Smith is now developing a series of training exercises for residents, attending physicians, and nurses, among others, where seasoned health care workers play out scripted clinical scenarios depicting microaggressions stemming from implicit bias—everyday putdowns, insults, or slights that minoritized populations face—while newer members of the medical establishment watch on. Health care workers witness microaggressions toward patients all the time, Smith says, but nobody teaches providers how to address the offenses. And so, after the simulations, Smith discusses the skits with these new practitioners and provides them with strategies to take action, such as documenting an inappropriate remark, when they see microaggressions in their workspace.

Sabin has also developed a 40-minute educational online course for medical school faculty across the country on how to manage bias. It includes a brief history of racism in medicine as well as advice on collecting data to identify inequities in care. Those who took the course not only increased their recognition of bias, but this awareness lasted for at least 1 year, Sabin and colleagues reported last year in a peer-reviewed publication. The participants credited the contents of the course for improving their teaching and their clinical practice. Sabin hopes the training can help these physicians be more thoughtful and mindful about avoiding stereotypes when filling patient charts, for example.

Niculescu and his team are addressing implicit bias from a different angle. They are trying to eliminate a central subjectivity in medical care by developing a blood test for biomarkers that reflect a person's level of pain. Objectively measuring pain "removes stigma, because people might be underappreciating your degree of pain or suffering," he says. "And blood biomarkers show that there is something biological going on. It's not something that you're just making up."

So far, the researchers have focused on identifying RNAs in blood that show the activity of specific genes and could indicate the level of someone's pain. Their preliminary results suggest the expression of the gene for a molecule called microfibril associated protein 3 (MFAP3) is lower when research participants are experiencing severe pain, and low activity in the gene is also a good predictor of future emergency room visits. MFAP3 hadn't previously been connected to pain and the group suggests it may normally have a pain suppression function.



However, validating biomarkers takes extensive replication studies in multiple, large populations. And even proven biomarker tests are usually not cheap, which makes them unlikely to be a reality anytime soon, especially for the marginalized populations that could benefit from their rollout. But Niculescu remains optimistic. "The medicine of the future hopefully will be equitable and not biased, and everybody will have access to these things," he predicts.

NONE OF THESE solutions on its own will eliminate implicit bias in medicine, researchers say. "We're not going to solve this tomorrow," says Charles Goldberg, an internal medicine physician at UCSD, who is also involved with the UnBIASED team. "But tomorrow is going to be better than today largely, and next month will be better than this month. And in a year, we're going to be moving along."

Getting buy-in from whole health care systems could accelerate the process. Recently, California, Michigan, Maryland, Minnesota, and Washington state passed legislation mandating implicit bias training for the medical professionals they license. And since June 2022, Massachusetts physicians are required to take implicit bias training to get a new license or get recertified to practice.

Although researchers see this as a good step, they worry mandated training will become a one-off box-checking exercise. Sustained implicit bias training for physicians should instead be the norm, some emphasize. Hospitals also need to monitor and collect data on health care outcomes for different groups in order to monitor equity, Sabin says. "You have to know where the disparities lie and then begin to work backwards from that."

It won't be easy, Hardeman says, noting that, at least in the United States, centuries of white supremacy and other forms of bigotry have resulted in deep-rooted stereotypes and other implicit biases. "Every single person should be thinking about doing this work," she says. "But if they're doing it within a system that hasn't addressed its own biases and racism, then it's not going to be fully effective." ■

ILLUSTRATIONS: THUMY PHAN

PHOTO: TODD EHLERS

The formation and erosion of mountains, such as the northern Pamir Mountains in Tajikistan, shown here, can have various impacts on the atmosphere and biosphere.

GEOLOGY

Landscapes through time

Changes to the landscape over millions of years are a driver of Earth system processes

By **Todd A. Ehlers**[1,2]

Landscapes are often considered by humans to be static. However, over time scales from the seconds it takes a sand grain to roll downhill to the tens of millions of years for a mountain to disappear by erosion, Earth's topography is highly dynamic. Understanding how landscapes change through time opens new research frontiers in quantifying the links between the geosphere, global and regional climate change, biodiversity, and biogeochemical cycles. On page 918 of this issue, Salles *et al.* (*1*) present a computer model reconstruction of global topography over the last 100 million years. The predicted fluxes of water, erosion, and sediment fit well with existing geologic observations and provide an understanding of Earth system interactions under conditions not present today.

Why is understanding how Earth's topography has evolved over millions of years important? In short, air, water, food, and energy from buried hydrocarbons result from interactions between Earth's surface, the biosphere, and the hydrologic cycle. The geologic record is rich with examples of "alternative Earth" states in which extreme conditions and environmental change are recorded. Understanding how past topography and sedimentary by-products changed through time is one important piece of the puzzle for understanding future Earth system responses, such as climate change. Salles *et al.* generate a landscape evolution model based on geomorphic transport laws for erosion and sediment transport, tectonic contributions to surface uplift that build topography, and paleoprecipitation changes that lower topography through erosion.

There are several ways topographic change over geologic time scales influences global and regional atmospheric circulation, composition, and the biosphere. These interactions occur over the thousands to millions of years required for plate tectonic processes to uplift mountains. When mountain topography exceeds 2 to 3 km above the neighboring lowlands, it obstructs atmospheric flow and influences precipitation and surface temperatures (*2*). Earth's most prominent mountain ranges, such as the Andes and the Himalayas, influence South American and Asian climatology (*3*, *4*), as well as global atmospheric phenomena such as the position of the Intertropical Convergence Zone (ITCZ) (*5*). In addition, the (abiotic) chemical weathering of silicate and carbonate rocks exposed in mountain ranges regulates global climate through the fixation and subsequent release of atmospheric carbon dioxide (CO_2). The efficiency with which this happens has recently been linked to erosion rates (*6*, *7*). Understanding the global pattern and temporal change of erosion across Earth's surface from studies such as that of Salles *et al.* is vital for future geochemical investigations that link physical erosion with chemical weathering to understand feedbacks between Earth's surface, atmosphere, and ocean chemistry.

The Earth's biota is closely linked to weathering, erosional, and depositional processes active on the surface (*8*–*10*). Biotic, chemical, and physical weathering processes break down rocks to form soils that sustain microbial communities, flora, and fauna. These interactions occur over the micrometers that mycorrhizal fungi supply plant nutrients from minerals, to the scale of mountain ranges where ecosystems and biomes influence the downslope transport and deposition of eroded sediment. Topography that forms in response to tectonic uplift and climate-driven surface processes over millions of years conspires with biota to control fundamental aspects of the Earth system, such as the production and storage of terrestrial and marine organic carbon in sediments that are shed from mountain ranges. The resulting fluxes of sediment and carbon control the long-term carbon cycle and the fossil energy resources on which humans depend. The landscape evolution model of Salles *et al.* predicts not only global patterns of erosion and the distribution of terrestrial and marine sediment deposits over the last 100 million years, but also the fluxes between them. Their results provide important predictions of temporal and spatial variations in the rates of deposition that will inform future studies of how global biogeochemical cycles are controlled.

Growing evidence suggests that topographic change during mountain building influences present-day biodiversity (*11*, *12*). Topography fosters biodiversity through its impact on local and regional climate gradients. The temperature and precipitation gradients resulting from topography lead to spatially variable environmental conditions that are conducive to diverse ecosystems. In addition, erosion associated with mountain building results in the exposure of different rock types. Spatial variations in rock types across mountains result in gradients in the structure and composition of soils and the nutrients available for ecosystems (*13*). Moreover, the effects of perturbations to Earth's climate (e.g., over glacial-interglacial cycles) on biota are buffered by topography through the creation of refugia at different elevations that species can retreat to and later expand from when more favorable conditions emerge. The study by Salles *et al.* does not explore biodiversity but rather provides fertile ground for future research using recent developments in evolutionary (*14*) and

[1]School of Geographical and Earth Sciences, University of Glasgow, Glasgow, UK. [2]Department of Geosciences, University of Tübingen, Tübingen, Germany. Email: todd.ehlers@glasgow.ac.uk

dynamic vegetation modeling (*15*).

These examples illustrate how topographic change and mountain building are coupled to tectonics, the biosphere, and atmospheric processes. However, process-based models such as those used by Salles *et al.* have their limits. They provide physics-based predictions of how eroded material from mountains is deposited in lowland and marine settings. But, simplifications are needed in the applied geomorphic transport laws to remain computationally feasible. As improved approaches emerge for formulating the physics and chemistry of erosion and deposition under different conditions, they need to be integrated into such models. Consideration of the glacial erosion effects on sediment accumulation is also necessary. These challenges require field and laboratory observations that can be upscaled to more extended time periods.

Understanding weathering, erosion, and deposition; topographic effects on biodiversity; and climate change during mountain building is still in its infancy. Future research efforts should focus on observationally driven parameterizations of these interactions. This will allow questions to be tackled, such as how do temporal and spatial variations in erosion and chemical weathering influence CO_2 drawdown? How did the coevolution of biota and landscapes influence mountain erosion, sediment accumulation, carbon fluxes, and storage? Although Salles *et al.* find that globally averaged sediment accumulation rates have remained somewhat constant over the past 100 million years, answering these questions (and others) requires detailed local and regional studies in which the signal of individual processes can be better understood and is not averaged at the global scale. This will involve advances in process-based landscape evolution models, investigations into the biotic effects on surface processes, and geochemical, stratigraphic, and geochronologic data sets that are suitable for evaluating model predictions. ■

REFERENCES AND NOTES

1. T. Salles *et al.*, *Science* **379**, 918 (2023).
2. G. Roe, *Annu. Rev. Earth Planet. Sci.* **33**, 645 (2005).
3. T. A. Ehlers, C. J. Poulsen, *Earth Planet. Sci. Lett.* **281**, 238 (2009).
4. S. Botsyun *et al.*, *Science* **363**, eaaq1436 (2019).
5. K. Takahashi, D. Battisti, *J. Clim.* **20**, 3434 (2007).
6. A. Bufe *et al.*, *Nat. Geosci.* **14**, 211 (2021).
7. R. G. Hilton, *Science* **379**, 329 (2023).
8. S. L. Brantley *et al.*, *Geobiology* **9**, 140 (2011).
9. J. Starke *et al.*, *Science* **367**, 1358 (2020).
10. M. Schaller, T. Ehlers, *Earth Surf. Dyn.* **10**, 131 (2022).
11. A. Antonelli *et al.*, *Nat. Geosci.* **11**, 718 (2018).
12. V. Sacek *et al.*, *Earth Planet. Sci. Lett.* **605**, 118033 (2023).
13. W. J. Hahm *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3338 (2014).
14. O. Hagen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2026347118 (2021).
15. C. Werner *et al.*, *Earth Surf. Dyn.* **6**, 829 (2018).

10.1126/science.adg5546

PHOTO: IRINA NAOUMOVA/ALAMY STOCK PHOTO

GENOMICS

Two domestications for grapes

Glacial cycles and wild adaptations shaped grape domestication and the rise of wine

By **Robin G. Allaby**

The domestication of plants that underpin the rise of human civilization is increasingly recognized as a complex interplay of processes across a culturally connected landscape. On page 892 of this issue, Dong *et al.* (*1*) reveal more of this complexity by reporting the unraveling of the evolutionary events that led to grape (*Vitis vinifera* ssp. *vinifera*) domestication. By incorporating the effects of glacial oscillations on biogeographical distributions of the wild progenitor (*Vitis vinifera* ssp. *sylvestris*) across Eurasia, they resolved two separate domestication processes from two distinct populations of *sylvestris* in the Near East and South Caucasus that were separated during the last glacial advance. They found that although the South Caucasus domestication is associated with early winemaking, the origin of wine in Western Europe is associated with cross-fertilization (introgression) between Western Europe's wild populations and domesticated grapes originating from the Near East that were initially used as food sources.

School of Life Sciences, Gibbet Hill Campus, University of Warwick, Warwick, UK. Email: r.g.allaby@warwick.ac.uk

The roots of domestication are frequently to be found deep in the Pleistocene, ending 11.5 thousand years ago (ka), where climate played a crucial role in determining human population densities and underlies the mosaic rates at which Neolithization occurred in the early Holocene, beginning 11.5 ka (*2*). Increasingly, ecological niche modeling is being used to track the past distributions of wild forms across these climatic shifts to help identify likely progenitor populations, for example in *Brassica* (*3*), as well the formation of subspecies after domestication, such as in rice (*4*). Incorporating this approach and using a titanic set of 2448 genomes from grapevine samples collected in 23 institutions across 16 nations around the world, Dong *et al.* establish that glacial episodes split *sylvestris* into eastern and western ecotypes—distinct and locally adapted varieties—around 500 ka. The last glacial advance saw the split of the eastern ecotype into two groups that each gave rise to a domestication process. It is this split of the eastern ecotypes that has not been previously resolved into two different domestications (*5*, *6*).

Despite being separated by more than 1000 km, the two domestication processes appear to have occurred contemporane-

dynamic vegetation modeling (*15*).

These examples illustrate how topographic change and mountain building are coupled to tectonics, the biosphere, and atmospheric processes. However, process-based models such as those used by Salles *et al.* have their limits. They provide physics-based predictions of how eroded material from mountains is deposited in lowland and marine settings. But, simplifications are needed in the applied geomorphic transport laws to remain computationally feasible. As improved approaches emerge for formulating the physics and chemistry of erosion and deposition under different conditions, they need to be integrated into such models. Consideration of the glacial erosion effects on sediment accumulation is also necessary. These challenges require field and laboratory observations that can be upscaled to more extended time periods.

Understanding weathering, erosion, and deposition; topographic effects on biodiversity; and climate change during mountain building is still in its infancy. Future research efforts should focus on observationally driven parameterizations of these interactions. This will allow questions to be tackled, such as how do temporal and spatial variations in erosion and chemical weathering influence CO_2 drawdown? How did the coevolution of biota and landscapes influence mountain erosion, sediment accumulation, carbon fluxes, and storage? Although Salles *et al.* find that globally averaged sediment accumulation rates have remained somewhat constant over the past 100 million years, answering these questions (and others) requires detailed local and regional studies in which the signal of individual processes can be better understood and is not averaged at the global scale. This will involve advances in process-based landscape evolution models, investigations into the biotic effects on surface processes, and geochemical, stratigraphic, and geochronologic data sets that are suitable for evaluating model predictions. ■

REFERENCES AND NOTES

1. T. Salles *et al.*, *Science* **379**, 918 (2023).
2. G. Roe, *Annu. Rev. Earth Planet. Sci.* **33**, 645 (2005).
3. T. A. Ehlers, C. J. Poulsen, *Earth Planet. Sci. Lett.* **281**, 238 (2009).
4. S. Botsyun *et al.*, *Science* **363**, eaaq1436 (2019).
5. K. Takahashi, D. Battisti, *J. Clim.* **20**, 3434 (2007).
6. A. Bufe *et al.*, *Nat. Geosci.* **14**, 211 (2021).
7. R. G. Hilton, *Science* **379**, 329 (2023).
8. S. L. Brantley *et al.*, *Geobiology* **9**, 140 (2011).
9. J. Starke *et al.*, *Science* **367**, 1358 (2020).
10. M. Schaller, T. Ehlers, *Earth Surf. Dyn.* **10**, 131 (2022).
11. A. Antonelli *et al.*, *Nat. Geosci.* **11**, 718 (2018).
12. V. Sacek *et al.*, *Earth Planet. Sci. Lett.* **605**, 118033 (2023).
13. W. J. Hahm *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3338 (2014).
14. O. Hagen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2026347118 (2021).
15. C. Werner *et al.*, *Earth Surf. Dyn.* **6**, 829 (2018).

10.1126/science.adg5546

PHOTO: IRINA NAOUMOVA/ALAMY STOCK PHOTO

GENOMICS

Two domestications for grapes

Glacial cycles and wild adaptations shaped grape domestication and the rise of wine

By **Robin G. Allaby**

The domestication of plants that underpin the rise of human civilization is increasingly recognized as a complex interplay of processes across a culturally connected landscape. On page 892 of this issue, Dong *et al.* (*1*) reveal more of this complexity by reporting the unraveling of the evolutionary events that led to grape (*Vitis vinifera* ssp. *vinifera*) domestication. By incorporating the effects of glacial oscillations on biogeographical distributions of the wild progenitor (*Vitis vinifera* ssp. *sylvestris*) across Eurasia, they resolved two separate domestication processes from two distinct populations of *sylvestris* in the Near East and South Caucasus that were separated during the last glacial advance. They found that although the South Caucasus domestication is associated with early winemaking, the origin of wine in Western Europe is associated with cross-fertilization (introgression) between Western Europe's wild populations and domesticated grapes originating from the Near East that were initially used as food sources.

School of Life Sciences, Gibbet Hill Campus, University of Warwick, Warwick, UK. Email: r.g.allaby@warwick.ac.uk

The roots of domestication are frequently to be found deep in the Pleistocene, ending 11.5 thousand years ago (ka), where climate played a crucial role in determining human population densities and underlies the mosaic rates at which Neolithization occurred in the early Holocene, beginning 11.5 ka (*2*). Increasingly, ecological niche modeling is being used to track the past distributions of wild forms across these climatic shifts to help identify likely progenitor populations, for example in *Brassica* (*3*), as well the formation of subspecies after domestication, such as in rice (*4*). Incorporating this approach and using a titanic set of 2448 genomes from grapevine samples collected in 23 institutions across 16 nations around the world, Dong *et al.* establish that glacial episodes split *sylvestris* into eastern and western ecotypes—distinct and locally adapted varieties—around 500 ka. The last glacial advance saw the split of the eastern ecotype into two groups that each gave rise to a domestication process. It is this split of the eastern ecotypes that has not been previously resolved into two different domestications (*5*, *6*).

Despite being separated by more than 1000 km, the two domestication processes appear to have occurred contemporane-

The transition from table grapes (left) to wine grapes (right) by wild introgression resulted in smaller berries with thicker skins, less sugar, and larger seeds.

ously with a high degree of shared signatures of selection on the same genes. Recent evidence demonstrates that long-distance networks of human communication and exchange over this distance occurred in the Epipalaeolithic, 10 to 20 ka (*7*, *8*). The findings of Dong *et al*. appear to show multiple emergences of domesticated forms occurring at a landscape level in which human communication would likely have been a key factor in promoting gene flow (*9*). To what extent the same domestication alleles might have been transported between different populations by humans or were present in the wild populations is a question that remains, and the answers will illuminate the role of human agency in grapevine domestication.

The South Caucasus domestication had limited spread and very little further influence, but the Near Eastern domestication came to dominate, establishing four major European cultivated grape clusters. Domestication is estimated to have occurred 11.5 ka, contemporaneous with the initial emergence of cereals, and the dates of the splits to form four European clusters match tantalizingly closely with the initial spread of the Neolithic into Europe. These dates are much earlier by several thousand years for both origin and spread than is expected from the archaeological evidence from domesticated seed morphology, which is distinct from wild seeds. The processes of selection that led to plant domestication can greatly predate the rise of morphological forms (*10*), so these earlier-than-expected dates may indicate exploitation of wild forms. Alternatively, although Dong *et al*. attempted to account for it, the long history of vegetative propagation (asexual reproduction) in grape cultivation (*11*) may be a confounding factor in divergence estimates by inflating generation time, which could explain the discrepancy with the archaeological record. To test between these alternatives will require direct investigation of the archaeological record using ancient DNA (archaeogenomics), as the authors also conclude.

The spread of domesticated plants from the Near East into new European environments is associated with a requirement for adaptation (*12*, *13*). Unlike cereals, grapes had wild populations in Europe from which they could obtain local adaptations. Gene flow from wild populations into domesticated forms can be hard to avoid and supplies both locally adapted variation and a resurgence of wild traits. Such adaptive introgression has also been noted in previous grape studies (*5*, *6*) and was recently reported in flax, which also originates from the Near East and has wild populations distributed throughout Europe (*14*). In both cases, the adaptive introgression is associated with a change in use. In flax, wild flowering-time genes introgressed and enabled adaptation to changes in daylength at higher latitudes. This was at the cost of seed size and oil content but also generated an architecture that was suitable for fiber production and may well have driven the textile revolution in central Europe, around 6 ka. In grapes, adaptations to the environment have been acquired that are associated with water stress and disease resistance (*1*, *5*). However, such introgressions also carried wild traits that compromise edibility. Compared with table grapes, wine grapes are smaller and thick skinned and have lower sugar content. These traits are more similar to those of wild grapes, which also makes them more suited for winemaking and less appealing for eating. That natural environmental adaptations underly the transition to wine raises key questions about the drivers behind such usage changes and to what extent they were forced by natural selection rather than by humans.

The enormous dataset produced by Dong *et al*. will provide insight into the finer points of grape evolution for some time to come. The increased resolution has pinpointed the lightening of berry color to some unknown genes close to the previously implicated *MybA* locus (*15*) and has suggested that the ancient Muscat flavor is unexpectedly rare possibly because of a pleiotropic constraint that prevents fixation. This study does not stray into the effects of structural variation in the genome, whereby many regions have been lost across various cultivar lineages. This has been shown to be key to the functional changes seen in domesticated grapes relative to their wild ancestors (*15*). The next big step will be to integrate these data into a structural landscape. ■

REFERENCES AND NOTES

1. Y. Dong, Y.-C. Lu, *Science* **379**, 892 (2023).
2. A. Palmisano *et al*., *Quat. Sci. Rev.* **252**, 106739 (2021).
3. M. E. Mabry *et al*., *Mol. Biol. Evol.* **38**, 4419 (2021).
4. R. M. Gutaker *et al*., *Nat. Plants* **6**, 492 (2020).
5. G. Magris *et al*., *Nat. Commun.* **12**, 7240 (2021).
6. S. Freitas *et al*., *Sci. Adv.* **7**, eabi8584 (2021).
7. E. Frahm, C. A. Tryon, *J. Archaeol. Sci. Rep.* **21**, 472 (2018).
8. L. A. Maher, *J. Archaeol. Method Theory* **26**, 998 (2019).
9. R. G. Allaby, C. J. Stevens, L. Kistler, D. Q. Fuller, *Trends Ecol. Evol.* **37**, 268 (2022).
10. R. G. Allaby *et al*., *Phil. Trans. R. Soc. B* **372**, 20160429 (2017).
11. J. Ramos-Madrigal *et al*., *Nat. Plants* **5**, 595 (2019).
12. A. Timpson *et al*., *J. Archaeol. Sci.* **52**, 549 (2014).
13. S. Shennan *et al*., *Nat. Commun.* **4**, 2486 (2013).
14. R. M. Gutaker *et al*., *Sci. Rep.* **9**, 976 (2019).
15. Y. Zhou *et al*., *Nat. Plants* **5**, 965 (2019).

10.1126/science.adg6617

BIOACOUSTICS

Voices in the ocean

Toothed whales evolved a third way of making sounds similar to that of land mammals and birds

By **Andrea Ravignani**[1,2] *and* **Christian T. Herbst**[3,4,5]

The ability of humans to sing and speak requires precise neural control of the larynx and other organs to produce sounds. This neural control is limited in most mammals (*1*). For animals that create complex sounds, less is known about how peripheral anatomical structures enable vocal feats (*2*). On page 928 of this issue, Madsen *et al*. (*3*) demonstrate that toothed whales, such as dolphins and killer whales, have a distinct nasal structure that produces diverse sounds in a broad frequency range that spans >4 orders of magnitude.

The findings of Madsen *et al*. stem from two long-standing strands of research: cetacean (toothed whales and baleen whales) communication and human voice science. For decades, studies of cetacean communication have relied on evidence from sound recordings complemented by postmortem anatomical investigation. However, cetaceans are large and patchily inhabit the ocean, so sampling sounds from specific individuals can be difficult and the rare postmortem samples cannot elucidate what happens "in action." As a result, finding the mechanism behind the sound production of toothed whales has proved elusive. In parallel, techniques have been developed to measure fine-grained dynamic parameters of the human vocal apparatus and map them to the sounds that are produced. Madsen *et al*. apply the methods used to study human voice to toothed whales to show that they blow air through their nasal passage and finely control it to produce diverse sounds.

This newly described voice production sys-

[1]Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands. [2]Center for Music in the Brain, Department of Clinical Medicine, Aarhus University and The Royal Academy of Music Aarhus/Aalborg, Aarhus, Denmark. [3]Department of Behavioural and Cognitive Biology, University of Vienna, Vienna, Austria. [4]Department of Vocal Studies, Mozarteum University Salzburg, Salzburg, Austria. [5]Janette Ogg Voice Research Center, Shenandoah Conservatory, Winchester, VA, USA. Email: andrea.ravignani@mpi.nl

The transition from table grapes (left) to wine grapes (right) by wild introgression resulted in smaller berries with thicker skins, less sugar, and larger seeds.

ously with a high degree of shared signatures of selection on the same genes. Recent evidence demonstrates that long-distance networks of human communication and exchange over this distance occurred in the Epipalaeolithic, 10 to 20 ka (*7*, *8*). The findings of Dong *et al*. appear to show multiple emergences of domesticated forms occurring at a landscape level in which human communication would likely have been a key factor in promoting gene flow (*9*). To what extent the same domestication alleles might have been transported between different populations by humans or were present in the wild populations is a question that remains, and the answers will illuminate the role of human agency in grapevine domestication.

The South Caucasus domestication had limited spread and very little further influence, but the Near Eastern domestication came to dominate, establishing four major European cultivated grape clusters. Domestication is estimated to have occurred 11.5 ka, contemporaneous with the initial emergence of cereals, and the dates of the splits to form four European clusters match tantalizingly closely with the initial spread of the Neolithic into Europe. These dates are much earlier by several thousand years for both origin and spread than is expected from the archaeological evidence from domesticated seed morphology, which is distinct from wild seeds. The processes of selection that led to plant domestication can greatly predate the rise of morphological forms (*10*), so these earlier-than-expected dates may indicate exploitation of wild forms. Alternatively, although Dong *et al*. attempted to account for it, the long history of vegetative propagation (asexual reproduction) in grape cultivation (*11*) may be a confounding factor in divergence estimates by inflating generation time, which could explain the discrepancy with the archaeological record. To test between these alternatives will require direct investigation of the archaeological record using ancient DNA (archaeogenomics), as the authors also conclude.

The spread of domesticated plants from the Near East into new European environments is associated with a requirement for adaptation (*12*, *13*). Unlike cereals, grapes had wild populations in Europe from which they could obtain local adaptations. Gene flow from wild populations into domesticated forms can be hard to avoid and supplies both locally adapted variation and a resurgence of wild traits. Such adaptive introgression has also been noted in previous grape studies (*5*, *6*) and was recently reported in flax, which also originates from the Near East and has wild populations distributed throughout Europe (*14*). In both cases, the adaptive introgression is associated with a change in use. In flax, wild flowering-time genes introgressed and enabled adaptation to changes in daylength at higher latitudes. This was at the cost of seed size and oil content but also generated an architecture that was suitable for fiber production and may well have driven the textile revolution in central Europe, around 6 ka. In grapes, adaptations to the environment have been acquired that are associated with water stress and disease resistance (*1*, *5*). However, such introgressions also carried wild traits that compromise edibility. Compared with table grapes, wine grapes are smaller and thick skinned and have lower sugar content. These traits are more similar to those of wild grapes, which also makes them more suited for winemaking and less appealing for eating. That natural environmental adaptations underly the transition to wine raises key questions about the drivers behind such usage changes and to what extent they were forced by natural selection rather than by humans.

The enormous dataset produced by Dong *et al*. will provide insight into the finer points of grape evolution for some time to come. The increased resolution has pinpointed the lightening of berry color to some unknown genes close to the previously implicated *MybA* locus (*15*) and has suggested that the ancient Muscat flavor is unexpectedly rare possibly because of a pleiotropic constraint that prevents fixation. This study does not stray into the effects of structural variation in the genome, whereby many regions have been lost across various cultivar lineages. This has been shown to be key to the functional changes seen in domesticated grapes relative to their wild ancestors (*15*). The next big step will be to integrate these data into a structural landscape. ■

REFERENCES AND NOTES

1. Y. Dong, Y.-C. Lu, *Science* **379**, 892 (2023).
2. A. Palmisano *et al*., *Quat. Sci. Rev.* **252**, 106739 (2021).
3. M. E. Mabry *et al*., *Mol. Biol. Evol.* **38**, 4419 (2021).
4. R. M. Gutaker *et al*., *Nat. Plants* **6**, 492 (2020).
5. G. Magris *et al*., *Nat. Commun.* **12**, 7240 (2021).
6. S. Freitas *et al*., *Sci. Adv.* **7**, eabi8584 (2021).
7. E. Frahm, C. A. Tryon, *J. Archaeol. Sci. Rep.* **21**, 472 (2018).
8. L. A. Maher, *J. Archaeol. Method Theory* **26**, 998 (2019).
9. R. G. Allaby, C. J. Stevens, L. Kistler, D. Q. Fuller, *Trends Ecol. Evol.* **37**, 268 (2022).
10. R. G. Allaby *et al*., *Phil. Trans. R. Soc. B* **372**, 20160429 (2017).
11. J. Ramos-Madrigal *et al*., *Nat. Plants* **5**, 595 (2019).
12. A. Timpson *et al*., *J. Archaeol. Sci.* **52**, 549 (2014).
13. S. Shennan *et al*., *Nat. Commun.* **4**, 2486 (2013).
14. R. M. Gutaker *et al*., *Sci. Rep.* **9**, 976 (2019).
15. Y. Zhou *et al*., *Nat. Plants* **5**, 965 (2019).

10.1126/science.adg6617

BIOACOUSTICS

Voices in the ocean

Toothed whales evolved a third way of making sounds similar to that of land mammals and birds

By **Andrea Ravignani**[1,2] *and* **Christian T. Herbst**[3,4,5]

The ability of humans to sing and speak requires precise neural control of the larynx and other organs to produce sounds. This neural control is limited in most mammals (*1*). For animals that create complex sounds, less is known about how peripheral anatomical structures enable vocal feats (*2*). On page 928 of this issue, Madsen *et al*. (*3*) demonstrate that toothed whales, such as dolphins and killer whales, have a distinct nasal structure that produces diverse sounds in a broad frequency range that spans >4 orders of magnitude.

The findings of Madsen *et al*. stem from two long-standing strands of research: cetacean (toothed whales and baleen whales) communication and human voice science. For decades, studies of cetacean communication have relied on evidence from sound recordings complemented by postmortem anatomical investigation. However, cetaceans are large and patchily inhabit the ocean, so sampling sounds from specific individuals can be difficult and the rare postmortem samples cannot elucidate what happens "in action." As a result, finding the mechanism behind the sound production of toothed whales has proved elusive. In parallel, techniques have been developed to measure fine-grained dynamic parameters of the human vocal apparatus and map them to the sounds that are produced. Madsen *et al*. apply the methods used to study human voice to toothed whales to show that they blow air through their nasal passage and finely control it to produce diverse sounds.

This newly described voice production sys-

[1]Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands. [2]Center for Music in the Brain, Department of Clinical Medicine, Aarhus University and The Royal Academy of Music Aarhus/Aalborg, Aarhus, Denmark. [3]Department of Behavioural and Cognitive Biology, University of Vienna, Vienna, Austria. [4]Department of Vocal Studies, Mozarteum University Salzburg, Salzburg, Austria. [5]Janette Ogg Voice Research Center, Shenandoah Conservatory, Winchester, VA, USA. Email: andrea.ravignani@mpi.nl

tem in the nasal passage of toothed whales has a striking similarity to what is known about vocalization in humans, terrestrial mammals, frogs, and birds, for which sound production typically occurs in the larynx or syrinx. An exhalatory aerodynamic force sets the laryngeal vocal folds or syringeal membranes into self-sustained oscillation with no need for recurring muscular contractions. The ensuing oscillatory tissue motion introduces cyclic variation into the exhalatory airstream. The resulting air-pressure variations are transmitted through the oral, pharyngeal, and nasal cavities (the "vocal tract") and are then emitted from the mouth and partly from the nose. This so-called myoelastic-aerodynamic (MEAD) principle (*4*) converts the exhalatory aerodynamic energy into sound.

Further parallels may exist between toothed whales and other mammals. The vocal sound generator in humans is a complex system of coupled subsystems (*5*), namely the oscillating tissue and its surrounding airspace (the trachea and vocal tract). Each of these physically interacting components produces its own set of biomechanically controlled resonance frequencies, which enables multiple types of oscillation for the sound generator (*6*). The distinct combinations of tissue vibratory modes that emerge establish different "voice registers." Alpine yodels and "voice cracks" in pubescent boys exemplify different vocal registers within one single bout of phonation.

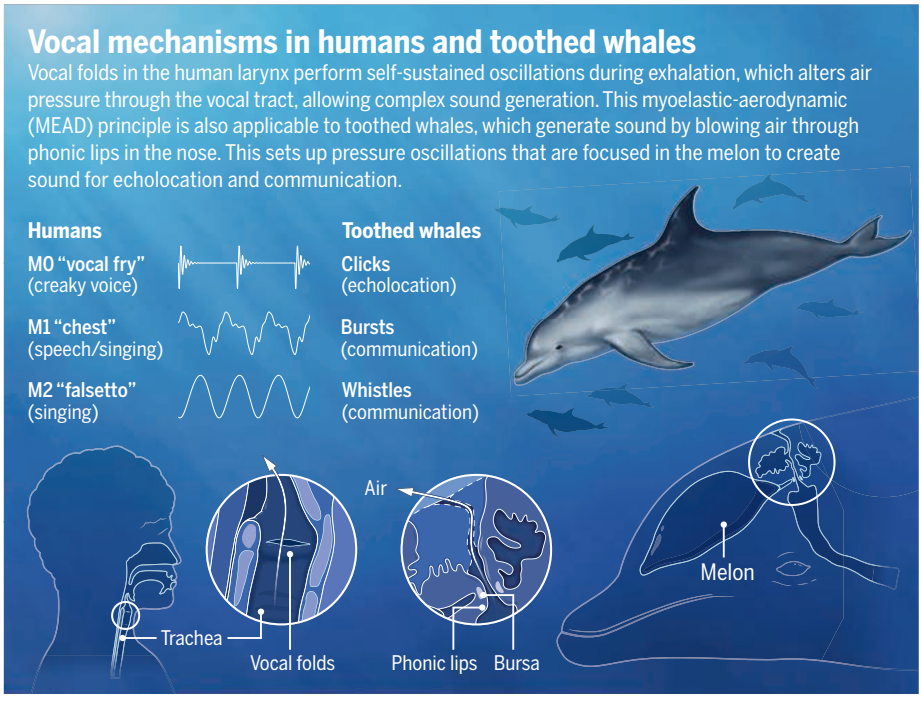
Three human voice registers (*7*) are physically analogous to those that Madsen *et al.* documented in toothed whales (see the figure). Mechanism M0, the "vocal fry," can be heard in "creaky voice," which often occurs in speech at the onset or offset of phrases. M1, the "chest register," is the stereotypical register in which humans speak. M2, the "falsetto register," is sometimes used in speech and also for exclamations and emotional responses. Voice registers are well-documented in human speech and singing, but finding these sound production mechanisms in toothed whales is unexpected.

Humans and almost all mammals make sounds by means of a laryngeal, self-sustained MEAD mechanism. Birds have a larynx but use their syrinx—a recently evolved organ whose MEAD mechanics resemble those of the mammalian larynx—to produce sound (*8*). Madsen *et al.* now show another versatile mechanism that toothed whales use to make sounds according to the MEAD principle: their nasal passage, whose mechanics resemble those of the larynx and the syrinx. Toothed whales not only produce clicks in the vocal fry (M0) register to echolocate but also social vocalizations in the chest (M1) and falsetto (M2) registers. The same vocal organ produces both pitched sounds in social communication and rhythmicity in echolocation clicks. This raises questions about perception: as the repetition rate of sounds becomes faster, humans stop perceiving them as rhythmic and start hearing a pitch. At what point does this transition happen in the hearing of toothed whales?

How did the similarities between toothed whales and all other MEAD-users emerge? Similar evolutionary pressures on two species may lead to functionally similar traits even when their last common ancestor lacked that trait. For example, doves and bats independently evolved wings; likewise, humans and toothed whales independently developed their main phonatory systems in the larynx and the nasal passage, piggybacking on different parts of the respiratory system. The physical mechanism of producing vocalizations through airflow-induced, self-sustained vibration of sound-generating tissue is so physically convenient that it keeps being repurposed, with at least three independent evolutionary strategies in mammals, birds, and toothed whales.

However, some mammals may also produce sounds without "normal" laryngeal phonation used by humans. For example, small koalas produce disproportionately "deep" sounds through a distinct vocal organ (*9*), but apparently also according to the MEAD principle. Conversely, cat purrs may arise through cyclic muscular contractions (*10*). Some seals haunt Antarctica with electronic-like vocalizations (*11*). *Indri indri* lemurs reach extreme sound intensities in their trumpet-like songs (*12*). The list continues with tongue-clicking bats (*13*) and whistling mice (*14*). Even more sound production mechanisms may be awaiting discovery.

Combined approaches are needed to understand the acoustic world around us. Mammalian bioacoustics excels at answering ethological questions, and songbird research has pinpointed the neural control mechanisms for song. Human voice research has described the physics and physiology of the sound production apparatus, thereby mapping potential pathologies and highlighting medical solutions. Singing voice research has described how the voice production system can be used artistically. The study of Madsen *et al.* adds to this important cross-fertilization of different scientific fields by revealing that toothed whales can display extraordinary vocal abilities while diving at 1000 meters and feasting on seafood. At least vocally, humans are not so special after all. ■



GRAPHIC: A. FISHER/*SCIENCE*

REFERENCES AND NOTES

1. K. Simonyan, *Curr. Opin. Neurobiol.* **28**, 15 (2014).
2. W. T. Fitch, B. de Boer, N. Mathur, A. A. Ghazanfar, *Sci. Adv.* **2**, e1600723 (2016).
3. P. T. Madsen, U. Siebert, C. P. H. Elemans, *Science* **379**, 928 (2023).
4. J. G. Švec, H. K. Schutte, C. J. Chen, I. R. Titze, *J. Voice* 10.1016/j.jvoice.2021.01.023 (2021).
5. C. T. Herbst, C. P. H. Elemans, I. T. Tokuda, V. Chatziioannou, J. G. Švec, *J. Voice* 10.1016/j.jvoice.2022.10.004 (2023).
6. Z. Zhang, *J. Acoust. Soc. Am.* **140**, 2614 (2016).
7. N. Henrich, *Logoped. Phoniatr. Vocol.* **31**, 3 (2006).
8. C. P. H. Elemans *et al.*, *Nat. Commun.* **6**, 8978 (2015).
9. B. D. Charlton *et al.*, *Curr. Biol.* **23**, R1035 (2013).
10. J. E. Remmers, H. Gautier, *Respir. Physiol.* **16**, 351 (1972).
11. C. Erbe *et al.*, *Acoust. Aust.* **45**, 179 (2017).
12. A. Zanoli *et al.*, *Am. J. Primatol.* **82**, e23132 (2020).
13. Y. Yovel, M. Geva-Sagiv, N. Ulanovsky, *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **197**, 515 (2011).
14. J. Håkansson *et al.*, *BMC Biol.* **20**, 3 (2022).

10.1126/science.adg5256

CHEMISTRY

Peptide barcodes meet drug discovery

Small-molecule libraries encoded by peptide tags may accelerate the search for therapeutics

By **Wolfgang Haap**

The success or failure of small-molecule drug discovery efforts strongly depends on the "hit-finding" approaches that are applied at the inception of the drug discovery program (*1*). High-throughput screening of compound collections is still the main strategy (*2*), but several other approaches have shown promise. These include screening virtual libraries using three-dimensional protein structure or ligand information (*3*), de novo design of ligands (*4*), screening fragment (very small molecule) libraries (*5*), screening (cyclic) peptide libraries (*6*), repurposing existing compounds, and screening DNA-encoded libraries (DELs) (*7*). On page 939 of this issue, Rössler *et al.* (*8*) reveal a new hit-finding method that uses peptide-encoded libraries (PELs), which are similar to DELs.

In PELs, solid-phase peptide and small-molecule syntheses are used to readily generate large libraries of bifunctional molecules that each consist of a peptide tethered to a small molecule through a cleavable linker. After cleavage from the solid phase, these libraries are incubated with an immobilized therapeutic protein of interest for affinity selection. To identify those molecules that bind to the target protein, the peptide is cleaved from the bifunctional molecule and sequenced using mass spectrometry technologies that are normally applied in proteomics research (such as nanoscale liquid chromatography—tandem mass spectrometry). On the basis of the sequence of the peptide, the chemical structure of the small-molecule ligand can be identified. This is because the single amino acids that are used for synthesis of the peptide directly encode the corresponding chemical building blocks that are used to synthesize the small molecules.

The advantages of PEL technology over DELs are manifold. Most notably, a PEL supports harsher and more diverse chemical reactions, including metal-catalyzed reactions and reactions that require strong acidic or basic conditions. This breadth enables the synthesis of a wider scope of drug-like molecules. Another advantage is the application of solid-phase synthesis for peptides and small molecules, which allows the use of excess reactants. This, in turn, supports a higher yield and purity of the final small molecules, which is expected to substantially improve the quality of the libraries. Changing the tagging moiety from four DNA bases to a peptide that contains 16 different amino acids enables a higher information capacity. Thus, in theory, even larger libraries of small molecules could be synthesized and encoded. If an eight-digit encoding string is used, then there are 16 amino acids (hexadecimal system) that can generate 4.3 billion possible codes. By contrast, there are only 56,535 possible codes using the four bases of DNA.

"[Peptide-encoded] libraries would be of high interest to drug discovery groups for screening against therapeutic protein targets for which no small-molecule ligands are yet known."

It is thought that the DNA tag of the DELs could interfere with targets that are per se DNA-binding, such as transcription factors or RNA. By contrast, libraries with peptide tags would potentially be better suited for screening against such targets because the amino acids used for the peptide synthesis are less likely to bind to those targets. To confirm that a hit identified in a DEL screen can actually bind to a target, the hit compound is synthesized without a DNA tag and then tested for its effect on biological activity. This can be tedious because, during DEL library synthesis, not every chemical reaction is successful. Occasionally, reaction by-products are the biologically active compounds, and it takes several investigations to determine this. The hit resynthesis that stems from a PEL can still be performed by solid-phase synthesis using the same conditions that were used to construct the library. This allows a more rapid synthesis and makes the identification of potential by-products easier.

Some challenges need to be overcome to fully exploit PEL technology. Peptide concentrations must be present in at least a 10 fM range to be detected by mass spectrometry. This affects the size of a PEL because, in contrast to a DNA tag, the peptide tag cannot be amplified. The screens of Rössler *et al.* were performed at ~1 nM concentration for each peptide-tagged compound. This means that a 100,000-membered library could be screened at a 100 μM library concentration. Because the peptide tags used by Rössler *et al.* were mostly hydrophobic, there is a certain risk of solubility problems and unspecific peptide aggregation of the library members, which could interfere with binding to a putative target and lead to screening artifacts.

The libraries generated by Rössler *et al.* were screened against the targets human carbonic anhydrase IX, the epigenetic reader bromodomain-containing protein 4 (BRD4), and the E3 ubiquitin ligase mouse double minute 2 homolog (MDM2). In all cases, several hits were identified that could serve as interesting starting points for further improvements of their potency and properties. PELs could potentially be enhanced by exploiting a wealth of already-established solid-phase organic chemistry reactions to generate new drug-like molecules in a chemical space that is not accessible by the DEL technology. Such libraries would be of high interest to drug discovery groups for screening against therapeutic protein targets for which no small-molecule ligands are yet known. ■

Research and Development Division, F. Hoffmann–La Roche AG, Basel, Switzerland.
Email: wolfgang.haap@roche.com

REFERENCES AND NOTES

1. D. G. Brown, J. Boström, *J. Med. Chem.* **61**, 9442 (2018).
2. P. S. Dragovich, W. Haap, M. M. Mulvihill, J.-M. Plancher, A. F. Stepan, *J. Med. Chem.* **65**, 3606 (2022).
3. F. Gentile *et al.*, *Nat. Protoc.* **17**, 672 (2022).
4. M. Skalic, J. Jiménez, D. Sabbadin, G. De Fabritiis, *J. Chem. Inf. Model.* **59**, 1205 (2019).
5. D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti, *Nat. Rev. Drug Discov.* **15**, 605 (2016).
6. C. Sohrabi, A. Foster, A. Tavassoli, *Nat. Rev. Chem.* **4**, 90 (2020).
7. R. A. Goodnow Jr., C. E. Dumelin, A. D. Keefe, *Nat. Rev. Drug Discov.* **16**, 131 (2017).
8. S. L. Rössler, N. M. Grob, S. L. Buchwald, B. L. Pentelute, *Science* **379**, 939 (2023).

ACKNOWLEDGMENTS

W.H. is an employee and stock owner of F. Hoffmann–La Roche Ltd.

10.1126/science.adg7484

POLICY FORUM

ARTIFICIAL INTELLIGENCE

The growing influence of industry in AI research

Industry is gaining control over the technology's future

By **Nur Ahmed**[1,2], **Muntasir Wahed**[3], **Neil C. Thompson**[1,2]

For decades, artificial intelligence (AI) research has coexisted in academia and industry, but the balance is tilting toward industry as deep learning, a data-and-compute-driven subfield of AI, has become the leading technology in the field. Industry's AI successes are easy to see on the news, but those headlines are the heralds of a much larger, more systematic shift as industry increasingly dominates the three key ingredients of modern AI research: computing power, large datasets, and highly skilled researchers. This domination of inputs is translating into AI research outcomes: Industry is becoming more influential in academic publications, cutting-edge models, and key benchmarks. And although these industry investments will benefit consumers, the accompanying research dominance should be a worry for policy-makers around the world because it means that public interest alternatives for important AI tools may become increasingly scarce.

INDUSTRY'S INPUT DOMINANCE

Industry has long had better access to large, economically valuable datasets (*1*) because their operations naturally produce data as they interact with large numbers of users and devices. For example, in 2020, WhatsApp users sent roughly 100 billion messages per day. Thus, it is unsurprising that most large data centers are owned and operated by industry [see supplementary materials (SM)]. In this article, we show that industry's dominance extends beyond data to the other key inputs of modern AI: talent and computing power.

Demand for AI talent has grown much more quickly than supply over the past decade (see SM), generating increased competition for AI talent. Across two different measures of talent, we see that industry is winning this contest. Data on North American universities (where we are able to get the best data) show that computer science PhD graduates specializing in AI are going to industry in unprecedented numbers (see the first figure). In 2004, only 21% of AI PhDs went to industry, but by 2020, almost 70% were. For comparison, this share of PhDs entering industry is already higher than in many areas of science and will likely soon pass the average across all areas of engineering (see SM). Computer science research faculty who specialize in AI have also been hired away from universities to work in industry. This hiring has risen eightfold since 2006, far faster than the overall increase in computer science research faculty (see the first figure). Between the PhD students and faculty leaving for industry, academic institutions are struggling to keep talent (*2*). This concern is not limited to US universities. In the UK, Abhinay Muthoo, Dean of Warwick University's King's Cross campus, said, "The top tech firms are sucking the juice from the universities" (*3*).

The computing power being used by academia and industry also shows a growing divide. In image classification, the computing power being used by industry is larger and has grown more rapidly than that used by academia or by industry-academia collaborations (see the first figure). Here, we proxy for the computing power used in a model with the number of parameters—both because the number of parameters is one of the key determinants of the computing power needed and because the deep learning scaling law literature has shown strong relationships between them. In 2021, industry models were 29 times bigger, on average, than academic models, highlighting the vast difference in computing power available to the two groups. This is not just a difference in approach but a shortfall in computing available to academics. For example, data from Canada's National Advanced Research Computing Platform reveals that academic demand for graphics processing units (GPUs; the most common chips used in AI) on their platform has increased 25-fold since 2013 (see SM), but supply has only been able to meet 20% of this demand in recent years.

Industry's ability to hire talent and harness greater computing power likely arises because of differences in spending. Although investments in AI have gone up substantially in both the public and private sectors, industry's investments are larger and growing faster (see SM). We compare industry with the major source of public-interest AI research: governments, which both fund their own research and are a key source of academic funding. In 2021, nondefense US government agencies allocated US$1.5 billion on AI. In that same year, the European Commission planned to spend €1 billion (US$1.2 billion). By contrast, globally, industry spent more than US$340 billion on AI in 2021, vastly outpacing public investment. As one example, in 2019 Google's parent company Alphabet spent US$1.5 billion on its subsidiary DeepMind, which is just one piece of its AI investment. In Europe, the disparity is smaller but is still present; AI Watch estimates that "the private and public sector account for 67% and 33% of the EU AI investments respectively" (*4*) (see SM). For comparison, in recent decades, research funding in the pharmaceutical industry has been split roughly evenly between the private sector and governments or nonprofits (see SM). An example of the scale of funding needed to pursue AI research comes from OpenAI, which began as a not-for-profit with the claim to be "unconstrained by a need to generate financial return" and aiming to "benefit humanity as a whole" (*5*). Four years later, OpenAI changed its status to a "capped for-profit organization" and announced that the change would allow them "to rapidly increase our investments in compute and talent" (*6*).

THE INCREASING DOMINANCE OF INDUSTRY IN AI RESEARCH

Industry's dominance of AI inputs is now manifesting in an increasing prominence in AI outcomes as well—in particular, in publishing, in creating the largest models, and in beating key benchmarks. Research papers with one or more industry co-authors grew from 22% of the presentations at leading AI conferences in 2000 to 38% in 2020 (see the second figure). Alternate definitions of what constitutes an industry paper yield substantially similar results (see SM). Industry's dominance is even more apparent in the largest AI models (*7*) and in benchmark performance. Industry's share of the biggest AI models has gone from 11% in 2010 to 96% in 2021 [see the second figure; data are from (*8*)]. We use model size as a proxy for the capabilities of large AI models, as is common in the literature. Model size is also often used as a proxy for computing power (see the first fig-

[1]Sloan School of Management, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. [2]Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA, USA. [3]Department of Computer Science, Virginia Tech, VA, USA. Email: nurahmed@mit.edu; neil_t@mit.edu

ure). This dual usage reflects how important compute is for predicting the performance of deep learning systems (*9*).

We investigate when academia, industry, or academia-industry collaborations led performance on AI benchmarks (see the second figure). When looking across these six benchmarks in image recognition, sentiment analysis, language modeling, semantic segmentation, object detection, and machine translation—as well as 14 more that cover areas such as robotics and common sense reasoning (see SM)—industry alone or in collaboration with universities had the leading model 62% of the time before 2017. Since 2020, that share has risen to 91% of the time. For example, sentiment analysis can be used to understand the emotional tone of written work. Until 2017, academia led this benchmark 77% of the time. But since 2020, industry alone or in collaboration has led 100% of the time. So whether measured by building state-of-the-art AI models (as measured by either size or benchmark performance) or by publishing in leading research outlets, our analysis shows industry's increasing prominence in AI outputs.

POLICY IMPLICATIONS

Industry's increasing investment in AI has the potential to provide substantial benefits to society through the commercializing of technology. Firms can create better products that benefit consumers [for example, machine translation benefits international trade (*10*)] and can streamline processes that drive down a firm's costs. Industry's investment in AI also produces tools that are valuable to the whole community (such as PyTorch and TensorFlow, which are widely used in academia), hardware that facilitates efficient training of deep-learning models [such as tensor processing units (TPUs)], and publicly accessible pretrained models (such as the Open Pretrained Transformer model by Meta).

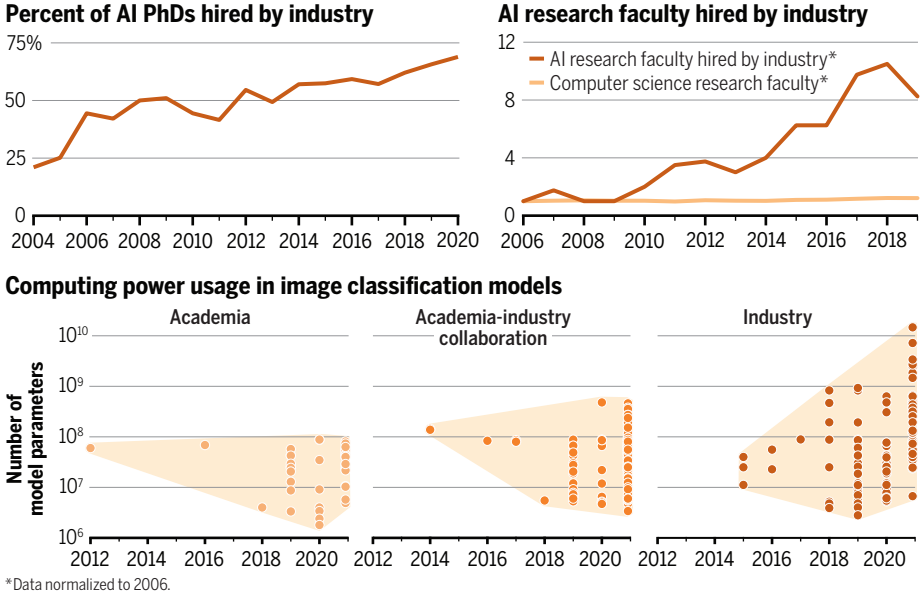
At the same time, the concentration of AI in industry is also worrisome. Industry's commercial motives push them to focus on topics that are profit oriented. Often such incentives yield outcomes in line with the public interest, but not always. Were all cutting-edge models from industry, situations would arise when no public-minded alternatives would exist. This possibility raises concerns akin to those about the pharmaceutical industry, where investment disproportionately neglects the needs of lower-income countries (*11*). Recent empirical work finds that "private sector AI researchers tend to specialise in data-hungry and computationally intensive deep learning methods" and that this is at the expense of "research involving other AI methods, research that considers the societal and ethical implications of AI, and applications in sectors like health" (*12*). These questions about the trajectory of AI and who controls it are also important for debates about job replacement and AI-induced inequality. Some researchers are concerned that we may be on a socially suboptimal trajectory (*13*) that focuses more on substituting human labor rather than augmenting human capabilities.

Even with a growing divide between industry and academia, one might imagine that the field could settle into a division of labor similar to that of other disciplines, in which basic research is primarily done in universities, and applied research and development is primarily done by industry. But in AI, such a clear divide does not exist; the same applied models used by industry are often those pushing the boundaries of basic research [a situation akin to what Donald E. Stokes referred to as "Pasteur's Quadrant" because of a similar overlap between applied and basic research in pasteurization (*14*)]. For example, transformers, a type of deep-learning architecture, were developed in 2017 by Google Brain researchers. Not only was this an important step forward in basic research, it was also applied almost immediately in models being used by industry. One benefit of this overlap is that it means that academic work can benefit industry directly (and industry has been supportive of efforts to increase public investment in AI). But this overlap also has a drawback: It means that industry domination of applied work also gives it power to shape the direction of basic research. Given how broadly AI tools could be applied across society, such a situation would hand a small number of technology firms an enormous amount of power over the direction of society. For many around the world, this concern is further heightened because these organizations are "foreign firms" to them. For example, the Future of Life Institute argues that "European companies are not developing general-purpose AI systems and are unlikely to start doing so anytime soon due to their relative competitive disadvantage vis-a-vis American and Chinese players" (*15*).

Even absent public alternatives to industry research, one might imagine that regulation, through auditing or external monitoring of industry AI, could be the solution. For example, in 2018 Joy Buolamwini, an academic, and Timnit Gebru, then a Microsoft employee, documented gender and racial biases in commercial face recognition systems (*16*). Establishing monitoring or auditing requirements (such as those in the Liability Rules for AI in Europe) can help mitigate these types of harms. However, if academics do not have access to industry AI systems, or the resources to develop their own competing models, their ability to interpret industry models or offer public-interest alternatives will be limited. This is both because academics would be unable to build the large models that seem to be needed for cutting-edge performance,

AI research inputs

(Top left) Percentage of US artificial intelligence (AI) PhDs hired by industry. (Top right) Growth of US university AI research faculty hired by industry, with a reference line for the total size of computer science research faculty. (Bottom) The total number of model parameters (a rough proxy for compute) for image recognition on ImageNet (see supplementary materials).



*Data normalized to 2006.

GRAPHIC: ADAPTED FROM AHMED *ET. AL* BY K. FRANKLIN/*SCIENCE*

but also because some useful capabilities of AI systems seem to be "emergent," meaning that systems only gain these capabilities once they are particularly large (*17*). Some negative characteristics of models also seem to scale with size [for example, toxicity in AI-generated language, and stereotyping (*7*)]. In either case, academics without access to sufficient resources would be unable to meaningfully contribute to these important areas.

Around the world, this concern about academia's resource disadvantage in AI research is being recognized, and policy responses are beginning to emerge. In the United States, the National AI Research Resource (NAIRR) task force (*18*) has proposed the creation of a public research cloud and public datasets. In Canada, the national Advanced Research Computing platform has been serving the country's academics and has been oversubscribed since its launch almost a decade ago. Chinese authorities have recently approved a "national computing power network system" (*19*) that will enable academics and others to access data and computing power. In Europe, similar initiatives have yet to emerge, although there is a clear recognition of the risk. As French president Emmanuel Macron said, "if you want to manage your own choice of society, your choice of civilization, you have to be able to be an acting part of this AI revolution" (*20*). For many countries, the scale needed for these types of investments may be daunting. In such cases, the key question for policy-makers will be whether they can pool sufficient resources with like-minded collaborators to reach the scale needed to create AI systems that reflect their own priorities.

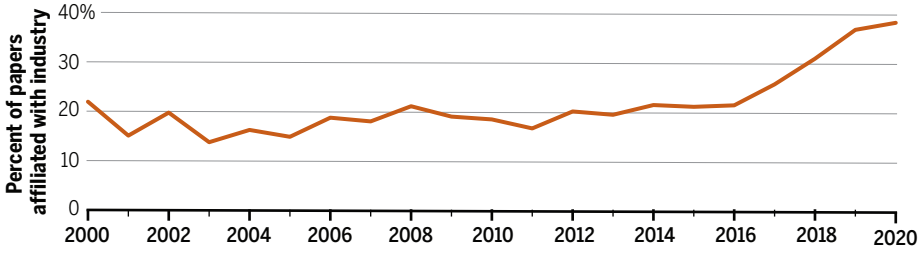
Computing power is not the only area in which remedies should be offered. Steps must also be taken for the other key inputs to AI. Building public datasets will be important but also a challenge because modern AI training datasets can be billions of documents. Of particular interest should be creating important datasets for which there are no immediate commercial interests. It is also important to provide the resources to keep top AI researchers in academia. For example, the Canada Research Chairs Program (CRCP), which provides salaries and research funds, has proven to be a successful means of attracting and retaining top talent in Canada.

For policy-makers working on this problem, the goal should not be that academia does a particular share of research. Instead, the goal should be to ensure the presence of sufficient capabilities to help audit or monitor industry models or to produce alternative models designed with the public interest in mind. With these capabilities, academics can continue to shape the frontier of modern AI research and benchmark what responsible AI should look like. Without these capabilities, important public interest AI work will be left behind. ■

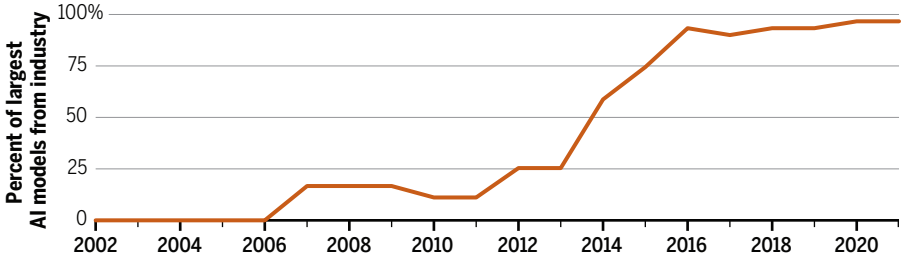
AI research outputs

(Top) The proportion of papers at leading AI conferences that have at least one industry co-author. (Middle) The fraction of the largest AI models that are from industry (3-year rolling average). (Bottom) Periods when the state-of-the-art model for leading AI benchmarks were from academia, industry, or collaborations (see supplementary materials).

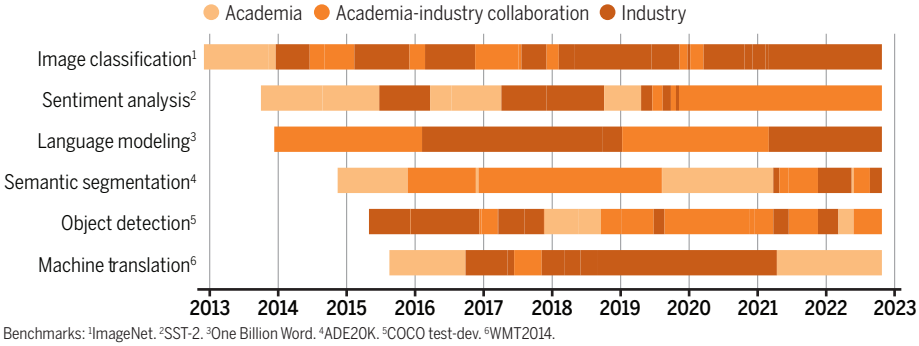
Publications by industry at leading AI conferences



Percent of the 10 biggest AI models that are from industry



Increasing domination of industry in AI benchmarks



Benchmarks: [1]ImageNet. [2]SST-2. [3]One Billion Word. [4]ADE20K. [5]COCO test-dev. [6]WMT2014.

REFERENCES AND NOTES

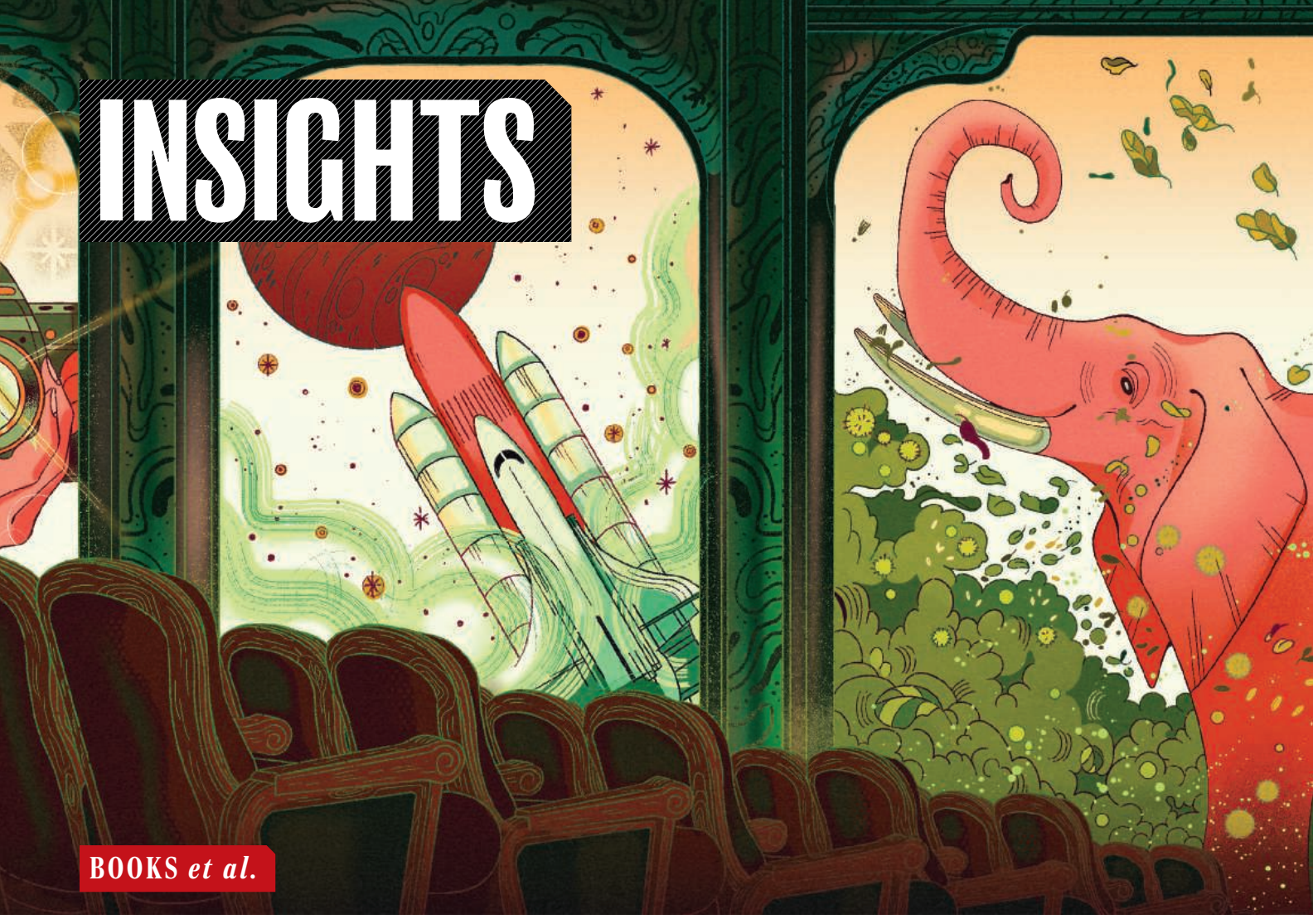
1. R. Shokri, V. Shmatikov, in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security* (ACM, 2015), pp. 1310–1321.
2. R. Jurowetzki, D. Hain, J. Mateos-Garcia, K. Stathoulopoulos, arXiv:2102.01648 [cs.CY] (2021).
3. "UK universities alarmed by poaching of top computer science brains," *Financial Times*, 9 May 2018.
4. T. Evas *et al.*, "AI Watch: Estimating AI investments in the European Union" (Publications Office of the European Union, 2022).
5. G. Brockman, I. Sutskever, OpenAI, "Introducing OpenAI," *OpenAI*, 11 December 2015.
6. G. Brockman, I. Sutskever, OpenAI, "OpenAI LP," *OpenAI*, 11 March 2019.
7. D. Ganguli *et al.*, in *2022 ACM Conference on Fairness, Accountability, and Transparency* (ACM, 2022), pp. 1747–1764.
8. J. Sevilla, L. Heim, M. Hobbhahn, T. Besiroglu, A. Ho, arXiv:2202.05924 [cs.LG] (2022).
9. N. C. Thompson, K. Greenewald, K. Lee, G. F. Manso, arXiv:2007.05558 [cs.LG] (2020).
10. E. Brynjolfsson, X. Hui, M. Liu, *Manage. Sci.* **65**, 5449 (2019).
11. P. Trouiller *et al.*, *Global Health* 267 (2017).
12. J. Klinger, J. Mateos-Garcia, K. A. Stathoulopoulos, *SSRN* 10.2139/ssrn.3698698 (2020).
13. E. Brynjolfsson, *Daedalus* **151**, 272 (2022).
14. D. E. Stokes, *Pasteur's Quadrant: Basic Science and Technological Innovation* (Brookings, 1997).
15. Future of Life Institute, "Emerging non-European monopolies in the global AI market" (Future of Life Institute, 2022); https://bit.ly/3k2ckD9.
16. J. Buolamwini, T. Gebru, in *Conference on Fairness, Accountability and Transparency* (PMLR, 2018), pp. 77–91.
17. J. Wei *et al.*, arXiv:2206.07682 [cs.CL] (2022).
18. D. E. Ho, J. King, R. C. Wald, C. Wan, "Building a national AI research resource," white paper (Stanford University Human-Centered Artificial Intelligence, 2021)
19. CAICT, "White paper on China's computing power development index" (CAICT, 2021); https://cset.georgetown.edu/wp-content/uploads/t0402_compute_white_paper_EN-2.pdf.
20. N. Berggruen, N. Gardels, *The Washington Post*, 27 September 2018.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade2420

10.1126/science.ade2420

GRAPHIC: ADAPTED FROM AHMED *ET AL.* BY K. FRANKLIN/*SCIENCE*



INSIGHTS

BOOKS *et al.*

REVIEW ROUNDUP

Science at Sundance 2023

A team of wildlife crime hunters work overtime to take down a dangerous ivory poaching ring. A couple tenderly navigates life with Alzheimer's disease. A space agency prepares for a manned mission to Mars. From a satirical glimpse into the future of human reproduction to a sobering look at the history of visual propaganda, a number of films featured at the 2023 Sundance Film Festival touched on topics and themes likely to be of interest to scientific audiences. Read on to see what our reviewers thought of seven of the films on offer this year. **—Valerie Thompson**

Poacher

Reviewed by **Vijaysree Venkatraman[1]**

A gunshot pierces the skull of an adult male elephant, a tusker, and it slowly slumps to the ground. This gruesome but gripping opening shot kicks off the engrossing episodic program *Poacher*, inspired by a true story, in which a motley team of wildlife crime fighters expose the largest ivory ring in Indian history.

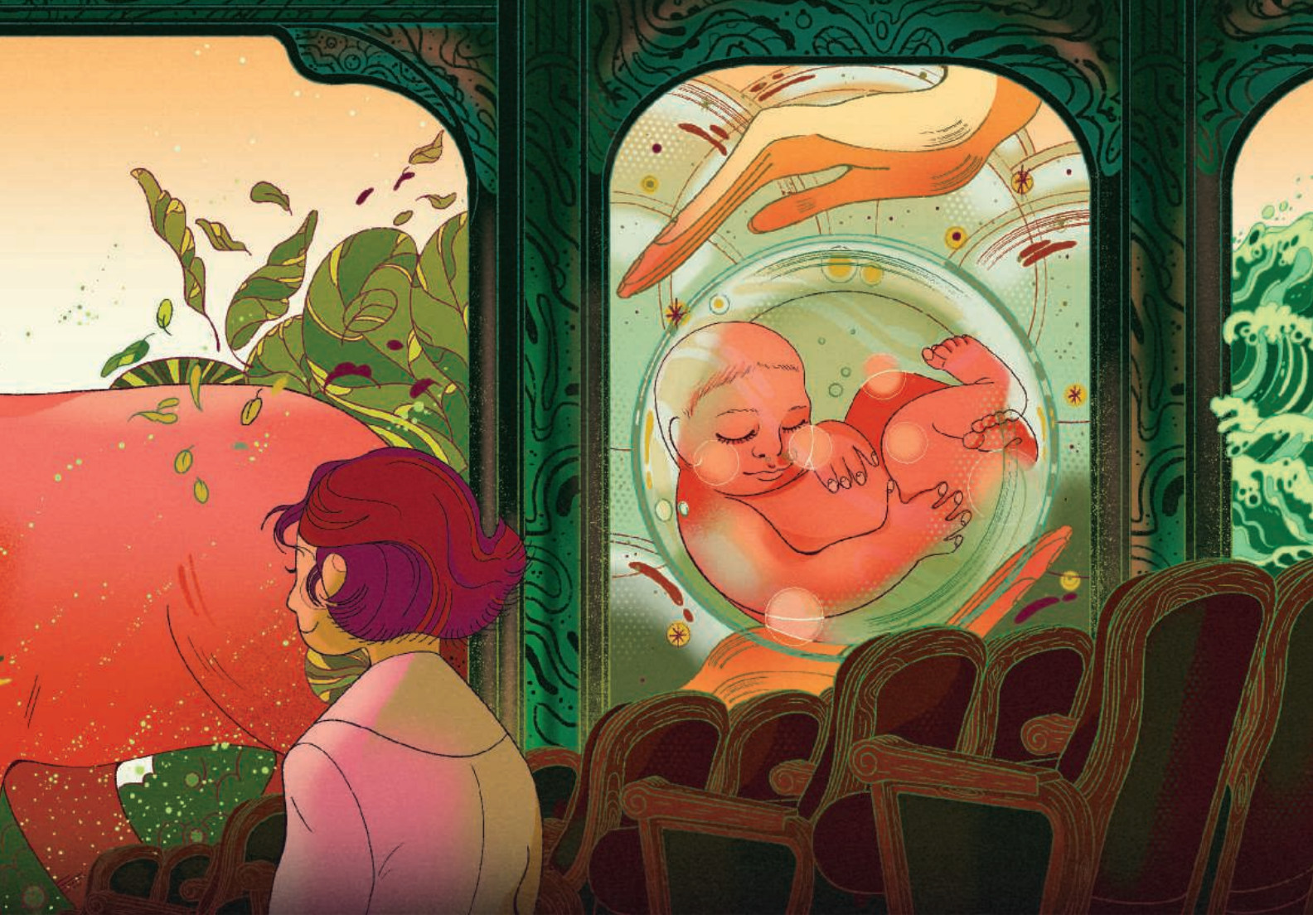
The series is set in 2015 in the jungles of Kerala, an Indian state nicknamed "God's own country," where, in 1994, authorities quashed an ivory-smuggling ring involving transnational crime syndicates. In the opening scene, a whistleblower comes forward to offer information on "Raaz," a dangerous elephant poacher the man claims is active in the area, but state officials are initially dismissive. Surveillance technology has become commonplace, and large-scale poaching is thought to be a thing of the past.

In this fast-paced investigative procedural—the first three episodes of which debuted at Sundance—narrative momentum is maintained without forfeiting character depth. India's religious diversity and the many languages spoken in the country are on full display: Viewers will hear Malayalam, English, and Hindi throughout the series.

The show's protagonists include computer programmer and snake expert Alan Joseph (played by Roshan Mathew), who builds a case against the poacher that "will live and die on data analysis," and forest officer Mala Jogi (Nimisha Sajayan), who leads raids into the hideouts of dangerous suspects and cleans up after the botched efforts of her colleagues. When the team eventually tracks down Raaz, viewers realize that this is just the beginning of a very complicated case.

Cleverly interspersed shots from the misty jungles suggest that the region's animals are

ILLUSTRATION: ISLENIA MIL



keeping a wary eye on the proceedings of the case too. If the elephants go, the jungle ecosystem will collapse, and Kerala will eventually be as polluted as the national capital, New Delhi, viewers are told. An aerial shot of the vehicle-clogged arteries of that megacity hints at what would be lost if this came to be.

Wildlife crime fighters are an overworked lot with little personal time. In *Poacher*, their triumphs and struggles are told with empathy. The dedication of these men and women to this dangerous work suggests that there is still hope for the future of wildlife on a planet where humans are now the top predators.

Poacher, *Richie Mehta, director*, QC Entertainment, 2022, 125 minutes.

Deep Rising

Reviewed by **Alison E. Barry**[2]

Using spectacular deep-sea cinematography, *Deep Rising* raises the alarm about the dangers of mining metallic nodules from the abyssal plains of the Pacific Ocean. The film weaves together, with partial success, several distinct story lines that suggest the complexity of its subject.

The film devotes much screen time to a start-up first known as DeepGreen, later renamed The Metals Company. Its CEO, Gerard Barron, tells audiences that nodules located on the seafloor are a rich source of the metals needed for batteries used in electric cars, thereby promoting green energy transition. According to Barron, the seabed is "the most desolate place on the planet," making development of this resource "clean" and far preferable to terrestrial strip mining. (Current scientific knowledge indicates a very high degree of biodiversity in the abyssal zone.)

Deep Rising shows both unsuccessful and successful deployments of a prototype nodule collector that resembles a cross between a combine and a vacuum. In its brief maiden voyage, it generates a huge cloud of fine sediment as it devours a small section of the seafloor. It is an impressive machine, but it is hard to discount concerns about disruption of the abyssal ecosystem.

Sandor Mulsow, a marine geologist and former director of environmental management at the International Seabed Authority (ISA)—an agency created by the United Nations Convention on the Law of the Sea (UNCLOS), which regulates the exploration and development of deep seabed mineral resources—frequently asserts in *Deep Rising* that the deep seabed is our "common heritage." This is not just his personal mission statement; he is quoting directly from UNCLOS, which was ratified in 1994 and is currently upheld by more than 160 member nations. (The United States signed the treaty in 1994, but the US Senate has never ratified it.) Tense scenes from ISA's hearings regarding proposed mining regulations reveal the conflicting motives and sentiments of various stakeholders.

The film uses Mulsow to bolster and lend credibility to its primary thesis: that mining to supply batteries for electric cars is unnecessary and will irretrievably disrupt fragile ecosystems to our detriment. But this is not as persuasive as the filmmakers evidently believed. Few facts

Alvy (Chiwetel Ejiofor) and Rachel (Emilia Clarke) embrace a tech-enabled pregnancy in *The Pod Generation*.

are mustered to support Mulsow's opinion about the availability of other battery technologies that do not require any kind of mining, even if we are willing to accept his message about environmental harm.

Ultimately, none of the subjects depicted in *Deep Rising* acknowledge the fundamental conflict between known benefits and unknown costs that besets our world. The film might have been more effective in advancing a meaningful discussion had the filmmakers chosen to focus on one of its many story lines or had explored the connections between them.

The views expressed are the author's own and are not intended to represent the opinion of DCP Midstream, LP, or its affiliates.

Deep Rising, *Matthieu Rytz, director*, Roco Films, 2022, 93 minutes.

The Pod Generation

Reviewed by **Nathaniel J. Dominy[3]**

Winner of the 2023 Alfred P. Sloan Feature Film Prize, *The Pod Generation*, written and directed by Sophie Barthes, envisions a believable future of artificial intelligence (AI)–mediated comforts. Guided by deep learning, "Elena"—an attentive and opinionated AI assistant—can brew a latte and print toast to perfection, all while offering wardrobe advice and monitoring one's "bliss index."

But the film is not so much about AI as it is a fable about Rachel (Emilia Clarke), a fast-rising tech executive, and her husband, Alvy (Chiwetel Ejiofor), a university botanist, and the challenges of pregnancy. The tension begins when Rachel's company, Pegazus, offers her a promotion with a corporate benefit: subsidized, priority access to its own subsidiary firm, The Womb Center, which produces and rents detachable wombs—gleaming egg-shaped pods—equipped with countless features to nurture, soothe, and stimulate the unborn baby.

Pegazus will even educate pod-born children in its own private schools, putting an insidious spin on its motives. Indeed, the honeycombed windows of The Womb Center are an inspired visual metaphor, suggesting a keen interest in the reproductive biology of eusocial insects, of controlling the growth and development of its own labor force.

For Rachel, the idea of a pod-incubated baby holds appeal for minimizing career disruptions, whereas Alvy, ever the nostalgist, prefers a natural process. Eventually the couple bring home a pod, with Alvy taking on most podcare responsibilities.

In one poignant scene, Alvy cradles the pod while watching Werner Herzog's *Encounters at the End of the World*. He is moved to tears by a "disoriented or deranged" Adélie penguin as it waddles "toward certain death." It is a tragicomic testament to the costs of social detachment, a theme mirrored in *The Pod Generation*. Increasingly unsettled by the intrusive technologies that surround them, the couple begins to wonder whether there is another way forward as their child's due date looms.

Through Alvy, the film touches on the complex pollination biology of figs, the symbiosis of lichens, and the "wood-wide web," giving viewers the sense that no organism is autonomous and independent, least of all humans. But these moments valorize Alvy at the expense of the concepts themselves. The biological consequences of integrating deep learning AI into the human holobiont, and the potential for host organism reversal, is a film for another time. Still, the egg-shaped artificial womb serves its purpose of startling the viewer, of satirizing and critiquing our overreliance on technology.

The Pod Generation, *Sophie Barthes, director*, MK2, 2022, 109 minutes.

The Longest Goodbye

Reviewed by **Vijaysree Venkatraman[1]**

In the next decade, NASA plans to send astronauts to Mars on a 3-year mission. The journey itself will take ~6 months each way. Whereas the various components of a spaceship can be tested under extreme conditions, the effects of prolonged social isolation on the crew members' emotional well-being remain unknown. And yet, how well the astronauts hold up mentally and emotionally in those cramped quarters could make or mar the mission. This simple but profound idea is elegantly explored in *The Longest Goodbye*.

The documentary features commentary from Al Holland, a NASA psychologist who is tasked with keeping space explorers mentally fit, as well as insightful interviews with astronaut Cady Coleman, who lived aboard the International Space Station (ISS) for 6 months from 2010 to 2011; Sukjin Han, a member of an Earthbound Mars simulation crew; and Kayla Barron, an astronaut currently in training for a potential Mars mission.

Archival video of Coleman's interactions with her family—which include a long-distance musical duet and a game of tic-

PHOTO: COURTESY OF SUNDANCE INSTITUTE

tac-toe—over a shaky internet connection during her 6-month stint at the ISS makes for heartwarming scenes. Coleman's son, Jamey, then in fourth grade, had a tough time with his mother's absence though. As he explains in the film, he always tried to put on a brave face for her.

"Crew members' connection with family is a critical piece of sustenance for them," Holland observes. Such connections are important during any long period of separation, learned Holland in 2010, when NASA was called in to help manage the mental health of 33 Chilean miners trapped underground. The documentary includes footage of the miners' 69-day ordeal and celebrated rescue, along with touching scenes of video calls with their families during their entrapment.

In a Mars expedition, astronauts will not be able to communicate with their families in real time, so experts are trying to come up with new strategies to counter homesickness. In the film, they discuss possible solutions, including virtual reality rendezvous with loved ones, AI-enabled companions, and even the possibility of inducing hibernation during the flight. A medical coma may spare the astronauts some angst en route but will likely lead to readjustment issues when they awaken, the experts concede.

Sometimes, the mission to Mars feels like too much to ask of any human for the sake of science. And yet the explorers who volunteer for such endeavors are often among the most eager participants. "If I could have spent another 6 months [on the ISS], I would have stayed in a minute," reveals Coleman in the closing moments of the film.

The Longest Goodbye, *Ido Mizrahy, director*, Autlook Filmsales, 2022, 87 minutes.

Is There Anybody Out There?

Reviewed by **Gabrielle Kardon[4]**

Scientists, like physicians, frequently approach disabilities as problems to be solved. *Is There Anybody Out There?* pushes back against this notion, offering viewers a glimpse into the full and complex lives of people with disabilities as they navigate the world and the worldviews of the able-bodied.

"Is there anybody out there with a body like mine?" is the question that motivates director Ella Glendining—who was born without hip joints and with short thigh bones—to document and explore her rare condition. Proximal femoral focal deficiency (PFFD) affects only 1 in every 200,000 children, few of whom have both legs affected, as Glendining does. Glendining has never met anyone who looks like herself and grabs a camera to document and explore her condition.

Glendining was born in Norfolk, UK, to supportive parents. However, when she started school, she realized she was not like her classmates. Archival footage of children with disabilities underscores how they are often made to feel different, with no potential for fulfilling lives. Glendining's interview with disability rights advocate Kevin Donnellon, a subject of one such film, compellingly shows how far from true this is.

In the film, Glendining discusses the challenges of finding work, where her job applications are denied because of her disability, and of pregnancy, where there are concerns about whether her PFFD will affect her ability to give birth. She is constantly exhausted from her daily interactions with people who cannot hide their shock when seeing her. "'It is what it is,' as they say on *Love Island*," she deadpans to the camera.

In her quest to find others who look like her, Glendining eventually meets Priscilla, Ricardo, and Charlie—all of whom have bilateral PFFD—and she feels less alone. Their disabled kinship is rejuvenating and joyous.

However, her efforts to find others with PFFD also reveal different attitudes toward her condition. Her parents took an "if it ain't broke, don't fix it" approach toward her body, yet others struggle to decide whether their children might be better off undergoing a complex series of surgeries to "fix" their limb abnormalities. Conversations with orthopedic surgeon Dror Paley and several parents facing the difficult decision of whether to subject their young children with PFFD to surgical interventions highlight the complex views held about bodies that differ from societal expectations.

Ella Glendining navigates ableist environments and worldviews in *Is There Anybody Out There?*

PHOTO: COURTESY OF SUNDANCE INSTITUTE. PHOTO BY DAVID MYERS

While she finds being disabled in an ableist world brutal, Glendining would not change a thing about herself. "I love and respect my body," she confidently confides. "Being this way is not the problem."

Is There Anybody Out There?, *Ella Glendining, director*, Hot Property Films Ltd, 2023, 87 minutes.

Fantastic Machine

Reviewed by **Nathaniel J. Dominy[3]**

Directed by Swedish visual sociologists Axel Danielson and Maximilien Van Aertryck, and winner of the World Cinema Documentary Special Jury Award for Creative Vision, *Fantastic Machine* is a sprawling history of camera technology and image-making. The film opens with the reactions of contemporary people experiencing a camera obscura, prompting one to exclaim, "Science, bitch!" It is a fitting expression of discovery and exhilaration, and it sets the tone for the first act.

The film then pivots to the origins of image permanence, beginning in 1827 with the earliest preserved photograph by Joseph Niépce. From there it unfolds swiftly and reverentially through the milestones of early photography, highlighting the contributions of Louis Daguerre (first person pictured), Eadweard Muybridge (first moving images), William Friese-Greene (first motion picture camera), and the Lumière brothers (first cinema with paying audiences). But this film is more than rote history.

It sharpens its focus with the reaction of King Edward VII to Georges Méliès's staged recreation of his coronation in 1902: "What a fantastic machine the camera is." This one-liner captures the love-hate thesis of *Fantastic Machine*—that humans have altered and often perverted the moving image to rewrite reality. Engaging but not revelatory, the film never acknowledges the irony of delivering its perspective through the very medium it criticizes. But perhaps the intent

[1]The reviewer is a freelance science journalist based in Boston, MA 02144, USA. Email: v.vijaysree@gmail.com [2]The reviewer is at DCP Midstream, LP, Denver, CO 80237, USA. Email: aebarry98@icloud.com [3]The reviewer is at the Department of Anthropology, Dartmouth College, Hanover, NH 03755, USA. Email: nathaniel.j.dominy@dartmouth.edu [4]The reviewer is at the Department of Human Genetics, University of Utah, Salt Lake City, UT 84112, USA. Email: gkardon@genetics.utah.edu

is to replicate the zeitgeist of our era, to produce an objective analysis that erodes the viewer's trust in objectivity.

This tension is palpable in the first sanguine words broadcast on Irish national television in 1961: "I hope this [television] service will provide for you all sources of recreation and pleasure, but also information, instruction, and knowledge." Although President Éamon de Valera added a prescient warning: "Like atomic energy, I am somewhat afraid. It can be used for incalculable good but also irreparable harm."

Fantastic Machine shines when it strives to instill image literacy. Vivid and fiercely edited are the scenes that juxtapose the work of Nazi propagandist Leni Riefenstahl—interviewed in 1993, she is practically giddy with her own skill at uplifting genocide—and British producer Sidney Bernstein, who sought to document the atrocities of the Holocaust. It is a meticulous and searing dissection of image-making in the service of truth and power.

The film ends with the "Pale Blue Dot" photo of Earth taken in 1990 by Voyager 1. It is a humbling image of introspection, not least because the Golden Record aboard the spacecraft contains 116 photos of humanity that pointedly exclude scenes of violence or suffering. An iconic instrument of science and verisimilitude, Voyager 1 is just another fantastic machine.

Fantastic Machine, *Axel Danielson and Maximilien Van Aertryck, directors*, See-Through Films, 2023, 88 minutes.

The Eternal Memory

Reviewed by **Gabrielle Kardon[4]**

The Eternal Memory, directed by Maite Alberdi, is a poignant and intimate portrait of a couple facing the challenges of Alzheimer's disease. The documentary follows Augusto Góngora, a prominent Chilean journalist who documented the brutal 17-year dictatorship of Augusto Pinochet in the 1970s and '80s, and Paulina Urrutia, a well-known actress and the first Chilean minister of culture.

In 2014 at the age of 62, Góngora was diagnosed with Alzheimer's, and Urrutia became his primary caregiver. When approached by Alberdi about a possible film project, Góngora convinced Urrutia to embrace the opportunity, saying, "I have no problem showing my fragility. I've made so many documentaries, why wouldn't I want to be filmed in this situation?"

The film records their journey over the past 4 years, a time in which Góngora's memory loss and dementia rapidly accelerated. Viewers see the couple's day-to-day activities—Urrutia reminding Góngora who he is and of their relationship, Urrutia gently washing and feeding Góngora, and the pair walking slowly around their neighborhood. Interspersed with such scenes is archival footage of Góngora's reporting, Urrutia's acting, and clips from old home movies.

Early in the film, Góngora has moments of lucidity, and his charisma and enjoyment of life are evident. He accompanies Urrutia to rehearsals and attends her performances with mild bemusement. However, toward the end, the ravages of the disease become more pronounced. Scenes filmed in the middle of the night by Urrutia, a measure necessitated by COVID-19 restrictions, intimately reveal his fragility and anguish. "Please help me," he pleads. "I can't go on like this."

The importance of memory and the tragedy of its loss are even more poignant given Góngora's work as a reporter. Góngora was a major contributor to *Chile: La Memoria Prohibida*, an important three-volume book documenting the events of the Pinochet regime and its consequences. Viewers watch as Urrutia reads Góngora's own words from the book: "Without memory we don't know who we are…we wander, confused, not knowing where to go…there is no identity."

The Eternal Memory, winner of the World Cinema Grand Jury Prize for Documentary at the 2023 Sundance Film Festival, is a profound and deeply intimate exploration of one couple's struggle with the losses imposed by Alzheimer's disease, but also the resilience of their love. Scientific audiences, normally focused on the etiology of Alzheimer's, will be reminded of the disease's devastating personal costs. ■

The Eternal Memory, *Maite Alberdi, director*, Micromundo/Fabula, 2023, 85 minutes.

10.1126/science.adg9997

Paulina Urrutia and Augusto Góngora grapple with the challenges of Alzheimer's disease in *The Eternal Memory*.

PHOTO: COURTESY OF SUNDANCE INSTITUTE

PHOTO: DIEGO CASTRO

Brazil could protect its free-flowing rivers by removing dams that are no longer operational, like the Pandeiros dam in Minas Gerais state.

Edited by **Jennifer Sills**

Conserving Brazil's free-flowing rivers

Brazil is home to nearly 34% of the world's remaining free-flowing waters (*1*). However, hydropower is the backbone of the Brazilian energy system, and the country's expanding dam infrastructure threatens its reserve of well-preserved rivers (*2*, *3*). Given the critical importance of connected rivers to ecosystems and people, Brazil's government should work to decrease the pace and minimize the impacts of dam construction.

Although President Luiz Inácio Lula da Silva's previous administration oversaw an expansion of hydropower, including controversial projects in the Amazon (*4*–*6*), he has pledged to embrace sustainability in his third term. Yet a market reserve created for small hydropower plants during the previous president's administration is expected to continue providing funds for existing and new small hydropower plant contracts (*7*). This investment in the construction of plants that use less than 50 MW will exacerbate river connectivity losses across the country (*3*). According to the national dam safety report, Brazil already has almost 23,000 existing dams, of which 2503 suffer from structural damage and poor maintenance (*8*).

Brazil's overinvestment in new construction and underinvestment in existing dams suggest a way to reconcile development and conservation. Retrofitting existing dams can help meet the country's energy needs without building new small hydropower plants (*9*), and dismantling useless and dangerous facilities can benefit people and fragmented aquatic ecosystems (*10*). Lula's third term represents an opportunity to adopt innovative river protection by advancing environmentally responsible policies and regulations on free-flowing rivers (*11*), hydropower site optimization (*12*), infrastructure retrofitting (*9*), and dam removal to restore river connectivity (*10*). Implementing such changes will require improved communication among policymakers, managers, and scientists.

Stephannie Fernandes[1]*, Thiago B. A. Couto[2], Manuel Ferreira[3], Paulo S. Pompeu[4], Simone Athayde[1], Elizabeth P. Anderson[1], Geraldo W. Fernandes[5,6]

[1]Florida International University, Miami, FL, USA. [2]University of Miami, Miami, FL, USA. [3]Universidade Federal de Goiás, Goiânia, GO, Brazil. [4]Universidade Federal de Lavras, Lavras, MG, Brazil. [5]Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil. [6]Knowledge Center for Biodiversity, Belo Horizonte, MG, Brazil.
*Corresponding author. Email: sdfernan@fiu.edu

REFERENCES AND NOTES

1. G. Grill *et al.*, *Nature* **569**, 215 (2019).
2. M. E. Ferreira *et al.*, *Water* **14**, 3762 (2022).
3. T. B. A. Couto, M. L. Messager, J. D. Olden, *Nat. Sustain.* **4**, 409 (2021).
4. P. M. Fearnside, Belo Monte, *Erde* **148**, 14 (2017).
5. T. B. A. Couto, J. D. Olden, *Front. Ecol. Environ.* **16**, 91 (2018).
6. S. Athayde *et al.*, *Curr. Opin. Environ. Sustain.* **37**, 50 (2019).
7. Gabinete de Transição Governamental, Relatório Final (2022); https://gabinetedatransicao.com.br/noticias/relatorio-final-do-gabinete-de-transicao-governamental/ [in Portuguese].
8. Agência Nacional de Águas e Saneamento Básico, Relatório de Segurança de Barragens (2021); https://www.snisb.gov.br/relatorio-anual-de-seguranca-de-barragem/2021/rsb-2021.pdf [in Portuguese].
9. K. Garrett *et al.*, *Environ. Res. Lett.* **16**, 114029 (2021).
10. J. R. O'Hanley *et al.*, *Landsc. Urban Plan.* **195**, 103725 (2020).
11. D. Perry *et al.*, *Sustainability* **13**, 1 (2021).
12. A. S. Flecker *et al.*, *Science* **375**, 753 (2022).

10.1126/science.adg9858

PFAS pollution threatens ecosystems worldwide

Toxic perfluoroalkyl and polyfluoroalkyl substances (PFASs) have been widely used for decades in textiles, food wrappings, flame retardants, water proofing, offshore industry, and cosmetics and are now ubiquitous in the environment (*1*). PFAS compounds are extremely resistant to biodegradation and persist for millennia, which complicates their management and cleanup (*2*). Their toxic properties make them carcinogenic, endocrine disruptive, and immunotoxic (*3*). In February, in response to PFASs' threat to ecosystems, biodiversity, food security, and drinking water (*4*, *5*), five EU countries proposed a complete phaseout of PFASs (*6*). Other countries should ban PFAS use as well.

PFASs are widely used in Europe and the United States. In Denmark, 500 tons

of two PFAS pesticides, diflufenican and fluopyram, have been sold for use on crop fields over the past decade (*4*). In the United States, common PFAS pesticides include malathion, an organophosphate insecticide (*7*, *8*). PFAS pesticides remain on crops, some of which are exported. As those products are used, consumed, and discarded around the world, human and environmental PFAS exposure grows (*9*).

More data are needed to respond to PFAS contamination. No programs meaningfully monitor the production, sources, fate, or transport of PFASs. Better environmental monitoring programs must be implemented globally, across aquatic and terrestrial ecosystems, to identify early warning signals and mitigate exposure, as has been done for the US Great Lakes ecosystems (*10*) and the Arctic (*11*). To minimize risk of PFAS residues in pesticides, pesticide manufactures, researchers, and policymakers must collaborate further (*12*).

Meanwhile, we must take immediate preventive action. The EU proposal is a good start—the European Chemicals Agency's scientific committees will begin a 6-month scientific evaluation in March (*9*), which could result in legislation. The United States and EU should use the momentum created by the evaluation process to initiate a complete global ban on PFASs and related compounds. Halting further pollution would support planetary health and UN Sustainable Development Goals. Policymakers could develop and enforce the ban through the Stockholm Convention and the related UN Environment Programme's chemical and waste initiatives (*9*).

Christian Sonne[1]*, Michael S. Bank[2,3], Bjørn M. Jenssen[4], Tomasz M. Cieseielski[4], Jörg Rinklebe[5], Su Shiung Lam[6,7], Martin Hansen[8], Rossana Bossi[8], Kim Gustavson[1], Rune Dietz[1]
[1]Aarhus University, Department of Ecoscience, Roskilde, Denmark. [2]Insititute of Marine Research, Bergen, Norway. [3]University of Massachusetts Amherst, Amherst, MA, USA. [4]Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway. [5]University of Wuppertal, School of Architecture and Civil Engineering, Wuppertal, Germany. [6]Universiti Malaysia Terengganu, Terengganu, Malaysia. [7]Saveetha Institute of Medical and Technical Sciences, Saveetha University, Center for Transdisciplinary Research, Chennai, India. [8]Aarhus University, Department of Environmental Science, Roskilde, Denmark.
*Corresponding author. Email: cs@ecos.au.dk

REFERENCES AND NOTES

1. J. Glüge *et al.*, *Environ. Sci. Process. Impacts* **22**, 2345 (2020).
2. J. W. Washington *et al.*, *Environ. Sci. Technol.* **43**, 6617 (2009).
3. P. Grandjean *et al.*, *PLOS One* **15**, e0244815 (2020).
4. "Farmers' use of PFAS pesticides can be a ticking time bomb," *Nyheder* (2023); https://nyheder.dk/landmaends-brug-af-pfas-pesticider-kan-vaere-en-tikkende-bombe/ [in Danish].
5. J. C. DeWitt *et al.*, *Toxicol. Pathol.* **40**, 300 (2012).
6. European Chemicals Agency (ECHA), "ECHA receives PFASs restriction proposal from five national authorities" (2023); https://echa.europa.eu/da/-/echa-receives-pfass-restriction-proposal-from-five-national-authorities.
7. S. Lasee *et al.*, *J. Hazard. Mater. Lett.* **3**, 100067 (2022).
8. M. G. Evich *et al.*, *Science* **375**, eabg9065 (2022).
9. R. Fuller *et al.*, *Lancet Planet. Health* **6**, e535 (2022).
10. US Environmental Protection Agency, Great Lakes Biology Monitoring Program (2022); https://www.epa.gov/great-lakes-monitoring/great-lakes-biology-monitoring-program.
11. AMAP and the Arctic Council (2023); https://www.amap.no/.
12. Z. Wang *et al.*, *Science* **371**, 774 (2021).

10.1126/science.adh0934

Potential of China's offshore wind energy

Floating offshore wind, which can operate in deep ocean waters, is a potential source for increasing renewable energy production (*1*). By 2035, 11 to 25% of all new offshore wind projects worldwide will feature floating equipment (*2*). However, this energy strategy faces technical, economic, and ecological challenges (*3*). By the end of 2021, only 17 floating offshore wind projects existed globally, with a cumulative installed capacity of only 142.37 MW (*4*). Countries planning massive investment in floating offshore wind farms, such as China, must ensure that the infrastructure is efficient and environmentally responsible.

In China, floating offshore wind is moving toward commercialization and large-scale development. In December 2022, China began building the world's largest commercial floating offshore wind farm off the coast of Hainan (*5*). The project will generate 11 times as much power as the current largest floating wind farm in Norway (*5*). Coastal provinces such as Shandong, Jiangsu, Fujian, and Guangdong are moving forward with plans to develop floating offshore wind installations by 2025. China's floating offshore wind is expected to reach 500 MW by 2026 (*6*).

China must make strategic decisions to ensure the efficient production of offshore wind energy. To avoid unnecessary construction, the country should analyze the optimal number and geographic configuration of wind turbines, as well as their contribution to decarbonized energy. Floating offshore wind infrastructure must seamlessly connect to the electrical grid to support the country's stability, reliability, and operational energy needs. Because the costs of floating technology are higher than those for fixed-bottom systems (*7*), sufficient capital investment across all phases will be crucial to the long-term success of China's floating offshore wind.

China must also protect the marine environment, given the potential for floating offshore wind to pose unique risks and outsized effects (*8*). The three-dimensional turbine motion increases the probability of turbine blade collisions (*9*). In addition, turbine infrastructure such as anchors and buried interarray cables could interfere with benthos and benthic habitats and disrupt biodiversity and seafloor integrity (*10*, *11*). In-depth monitoring and assessment of the cumulative impacts—both positive and negative—of floating offshore wind equipment will be essential. If necessary, the country should adopt new technologies and adaptive management practices to minimize adverse environmental impacts (*12*). Failing to do so could turn China's floating offshore wind farms into another example of extractivist development.

Siyou Xia, Yu Yang*, Yi Liu
Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing, China and College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China.
*Corresponding author.
Email: yangyu@igsnrr.ac.cn

REFERENCES AND NOTES

1. P. Rosa-Aquino, "Floating wind turbines could open up vast ocean tracts for renewable power," *The Guardian* (2021).
2. R. Wiser *et al.*, *Nat. Energ.* **6**, 555 (2021).
3. I. Galparsoro *et al.*, *npj Ocean Sustain.* **1**, 1 (2022).
4. L. Zhao, *Wind Energ.* **5**, 54 (2022) [In Chinese].
5. D. Cheong, "China building world's biggest offshore wind farm as renewable energy growth gains momentum," *The Straits Times/Asia News Network* (2023); https://www.straitstimes.com/asia/east-asia/china-building-world-s-biggest-offshore-wind-farm-as-renewable-energy-growth-gains-momentum.
6. "China's floating offshore wind power is expected to reach 500 MW in 2026," *Seetao* (2022); https://www.seetao.com/details/180755.html.
7. P. Beiter, W. Musial, L. Kilcher, M. Maness, A. Smith, "An Assessment of the Economic Potential of Offshore Wind in the United States from 2015 to 2030," *Tech. Rep.* (NREL/TP-6A20-67675, 2017).
8. A. F. Johnson, C. L. Dawson, M. G. Conners, C. C. Locke, S. M. Maxwell, *Science* **376**, 6591 (2022).
9. H. Bailey, K. L. Brookes, P. M. Thompson, *Aquat. Biosyst.* **10**, 8 (2014).
10. H. K. Farr *et al.*, *Ocean Coast. Manage.* **207**, 105611 (2021).
11. A. E. Copping, L. G. Hemery, "OES-Environmental 2020 State of the Science Report: Environmental Effects of Marine Renewable Energy Development Around the World," Report for Ocean Energy Systems (OES) (2020).
12. A. Copping, V. Gartman, R. May, F. Bennet, in *Wind Energy and Wildlife Impacts: Balancing Energy Sustainability with Wildlife Conservation*, R. Bispo, J. Bernardino, H. Coelho, J. Lino Costa, Eds. (Springer International Publishing, 2019), pp. 1–25.

10.1126/science.adh0511

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

For some fish species, adding shallow water habitat to lakes is a more effective management strategy than stocking fish.

PHOTO: FLORIAN MÖLLERS/AVN

CONSERVATION ECOLOGY

Adding habitat beats adding fish

Conservation and management efforts are often focused on protecting individual species. Alternatively, management can target restoring ecosystem processes or broader habitats. Such ecosystem-based management practices have limited support because of their high cost and unknown efficacy compared with species-focused approaches. Radinger *et al*. tested the effects of two habitat-based interventions, creating shallow zones and adding dead wood, to the more common approach of stocking fish species. Across their 20 experimental lakes, the authors found that fish stocking was ineffectual, whereas shallow zone creation increased target fish abundance, especially that of juvenile fish. This study demonstrates the potential for ecosystem-based management to meet conservation goals. —BEL *Science*, adf0895, this issue p. 946

CORRELATED ELECTRONS

Sluggish valence fluctuations

The strange metal phase, which can form in some strongly correlated materials, is characterized by the breakdown of the usual charge transport laws. Understanding charge dynamics in this exotic phase is, however, hampered by the lack of suitable probes. Kobayashi *et al*. used synchrotron radiation–based Mössbauer spectroscopy to study these dynamics in the strange metal phase of the heavy fermion metal β-$YbAlB_4$. The researchers observed a splitting of the Mössbauer absorption peak, which they attributed to unusually slow fluctuations of the ytterbium ion valence in this material. —JS

Science, abc4787, this issue p. 908

COMPUTATIONAL BIOLOGY

Designer PPIs

Efficient signal transduction in cells depends on selective protein-protein interactions (PPIs). In nature, new PPIs arise from gene duplication and mutation events that endow gene products with new interaction partners and isolate them from those of their ancestors (orthogonal interactions). Malinverni and Babu devised an algorithm relying on sequence analysis that simulated natural gene duplication and mutational events to design orthogonal PPIs. The authors validated this approach by showing that it recapitulated experimentally determined PPIs in a bacterial two-component system. —JFF

Sci. Signal. **16**, eabm4484 (2023).

WILDFIRE EMISSIONS

Emission emergency

Carbon dioxide emissions from boreal forest fires have been increasing since at least the year 2000, reaching a new high in 2021, Zheng *et al*. report. Although boreal fires typically produce about 10% of global carbon dioxide emissions from wildfires, in 2021 they produced nearly one quarter of the total. This abnormally high total resulted from the concurrence of water deficits in North America and Eurasia, which was an unusual situation. The increasing number of extreme wildfires that is accompanying global warming presents a real challenge to global climate change mitigation efforts. —HJS

Science, ade0805, this issue p. 912

PLANT SCIENCE

Engineered immune receptors

The specificity of plant innate immune receptors cannot rapidly change in response to new pathogens. To increase the diversity of pathogens that plants can sense, Kourelis *et al*. exploited the specificity and versatility of mammalian antibodies. The authors modified rice-derived receptors

that normally perceive fungal pathogen effectors. Part of the receptor protein was swapped with camelid antibody fragments that recognize fluorescent proteins. Plants expressing these chimeric proteins were able to mount an immune response when fluorescent proteins were introduced. The work provides a way forward to rapidly and specifically modify susceptible crop species to provide them with pathogen resistance. —MRS

Science, abn4116, this issue p. 934

ACTIVE MATTER

Getting in your lane

Relatively sparse active systems, ranging in scale from atoms to pedestrian traffic, tend to form lanes when forced to cross from two different directions. Bacik *et al*. derived a new formulation for the angles of those lanes based on simple rules in the case of two intersecting groups of pedestrians. In an experimental arena, directing people to walk toward each other with simple rules, such as moving to the right to avoid a collision, generated a number of interesting laning scenarios. Backed by numerical simulations, the authors developed a general model for determining when laning occurs and how it will look depending on the microscopic variables of a system. —BG

Science, add8091, this issue p. 923

Models help explain how lanes emerge in active systems such as pedestrians walking in opposing directions, as in this busy intersection in Hong Kong.

MUSCULAR DYSTROPHY

Activating adenylyl cyclase to treat DMD

Duchenne muscular dystrophy (DMD) is characterized by repeated contraction-induced muscle injury that eventually leads to failure of skeletal muscle regeneration. Extraocular muscles (EOMs) are generally spared in DMD, suggesting the presence of a potential compensatory mechanism. Taglietti *et al.* generated a rat model of DMD and found that EOMs had increased thyroid stimulating hormone receptor (TSHR), resulting in decreased muscle stem cell senescence. Forskolin, an adenylyl cyclase activator, stimulated TSHR signaling and led to improved proliferation and reduced senescence of skeletal muscle stem cells, resulting in improved functional performance in affected rats. These findings suggest that activation of TSHR signaling may have therapeutic value in patients with DMD. —MN

Sci. Transl. Med. **15**, eadd5275 (2023).

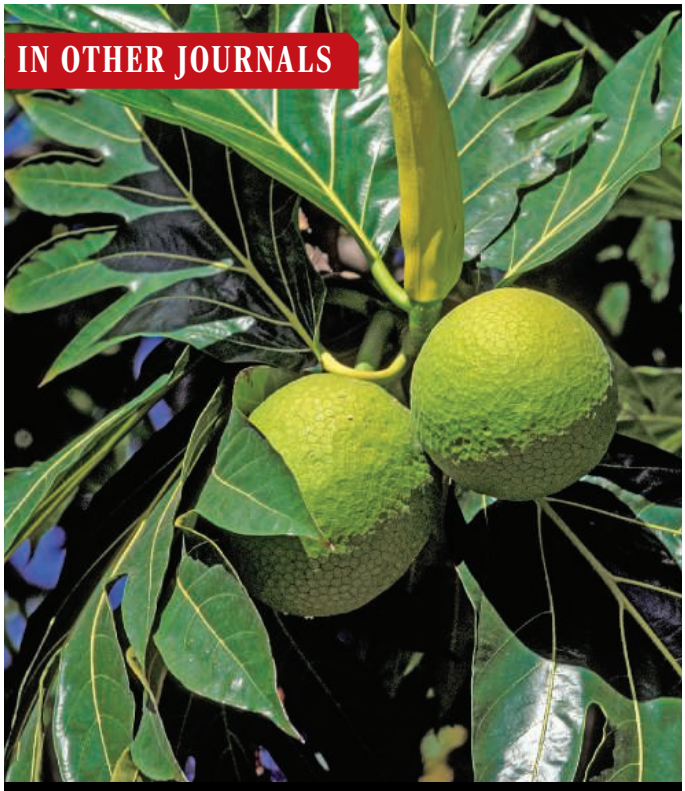
HEMATOPOIESIS

Trimming microRNA for blood development

Mutations in the 3′ to 5′ RNA exonuclease USB1 cause a pediatric disease with defects in the production of blood cells, although the underlying cause of this syndrome is unknown. Jeong *et al*. determined that USB1 removes extra adenosines from the 3’ end of microRNAs, which if not removed, promote degradation of microRNAs that are necessary for blood development. Blocking the enzyme that adds extra adenosines to microRNAs also restores the production of blood cells in mutant settings. This work determines the molecular basis of disease, identifies new roles for USB1, and suggests a possible therapeutic intervention for patients. —DJ

Science, abj8379, this issue p. 901

IN OTHER JOURNALS



PLANT BIOGEOGRAPHY

The terrible history of breadfruit

In 1789, Captain William Bligh’s attempt to export breadfruit plants from Tahiti to the Caribbean was foiled by mutiny. By 1793, breadfruit was successfully introduced to St. Vincent, from where it was distributed across the Caribbean as cheap food for enslaved people. Many island communities exposed to climate change now rely on breadfruit as a staple, so it is important to understand its genetics to enable appropriate management. Audi *et al.* combined genetic sequence analysis of 238 individual breadfruit with anatomical characterization, historical records, and local knowledge. The authors discerned five major Caribbean lineages of breadfruit, which match voyage records for the five types originally introduced from East Polynesia. Two Caribbean cultivars from St. Vincent and the Grenadines were confidently found to match extant Tahitian and Timorese cultivars genetically, but several other possible matches were identified. —MRS *Curr. Biol.* **33**, 287 (2023).

Breadfruit is a starchy food crop originally brought to the Caribbean to cheaply feed enslaved people.

ANCIENT DNA

Unearthing history in Charleston

Ancient DNA can be a powerful tool in uncovering the population history of individuals whose origins have been lost or intentionally obscured, as in the case of enslaved peoples. Working with community stakeholders, Fleskes *et al.* examined genomic and mitochondrial DNA from 36 individuals from the 18th century whose interment was discovered during construction in South Carolina. No known historical record of

PHOTOS: (LEFT TO RIGHT) WOLFGANG KAEHLER/LIGHTROCKET VIA GETTY IMAGES

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

PLANT GENETICS

Origins and domestication of grapevines

Humans have extensively shaped the organisms around us through domestication. Although wine and table grapes have been important culturally for thousands of years, their origin has been difficult to pinpoint because of uneven sampling of modern cultivars. Dong *et al*. analyzed genetic data from about 3500 cultivated and wild grape varieties from around the world (see the Perspective by Allaby). The results of their analysis reveal the effects of climate on historic population sizes, suggest concurrent domestications of wine and table grapes, and identify variants associated with domestication traits such as berry color and palatability. These results increase our understanding of how humans and the environment shaped this domesticated crop. —CNS

Science, add8655, this issue p. 892; see also adg6617, p. 880

LANDSCAPE DYNAMICS

Shaping Earth's surface

How has sediment transfer and accumulation from lands to oceans affected Earth's geomorphology? Salles *et al*. present results from a high-resolution model of surface physiography, validated by independent observations from the geological record, which simulates landscape evolution caused by erosion and deposition over the past 100 million years (see the Perspective by Ehlers). Their results should help us better understand the apparent contradiction between the observed Late Cenozoic pulse in marine sedimentation and the constancy of the global weathering flux. —HJS

Science, add2541, this issue p. 918; see also adg5546, p. 879

VOCALIZATION

Secrets of whale vocal anatomy

Odontocete ("toothed") whales are well known for using echolocation to forage underwater, but they also produce a wide array of sounds used for social communication. Precisely how all of these sounds are produced was characterized by Madsen *et al*. in living animals (see the Perspective by Ravignani and Herbst). The authors found that the wide array of sounds are produced through the nasal passages but in a way analogous to laryngeal and syringeal sound production. Using the nasal passages facilitates complex sound production at depth, where laryngeal sound would be hampered by pressure. Furthermore, they found that odontocetes use different vocal registers, such as those we associate with falsetto or vocal fry in humans, to convey information. —SNV

Science, adc9570, this issue p. 928; see also adg5256, p. 881

DRUG DISCOVERY

Peptide tags for small molecules

During the early stages of drug discovery, chemists often expose target proteins to vast libraries of small molecules in the hope of finding one that binds tightly. Tagging the molecules with small fragments of DNA has proven a convenient means of interpreting the screen at concentrations where it would otherwise be hard to determine which of them were bound. However, nucleotide structure constrains the chemistry applicable to making the drug candidates. Rössler *et al*. now showcase an alternative method in which the tags consist of peptides in place of oligonucleotides, expanding the scope of compatible chemistry (see the Perspective by Haap). —JSY

Science, adf1354, this issue p. 939; see also adg7484, p. 883

DENDRITIC CELLS

Building out the DC family tree

Analysis of dendritic cell (DC) subsets has demonstrated extensive phenotypic and functional heterogeneity. Identifying the progenitor cell populations that serve as precursors for the three main subsets of conventional DCs (cDCs) has been challenging. Rodrigues *et al*. used lineage tracing in mouse models and in vivo transfer studies to analyze the differentiation of a DC subset called "pDC-like cells" with morphological features of plasmacytoid DCs (pDCs) but other features suggestive of cDCs. These pDC-like cells were identified as precursors of a substantial fraction of the cDC2 subset through a differentiation process dependent on the KLF4 transcription factor. These studies revealed a key developmental pathway used by an unusual set of DC precursor cells with features conserved between mice and humans. —IRW

Sci. Immunol. **8**, eadd4132 (2023).

that normally perceive fungal pathogen effectors. Part of the receptor protein was swapped with camelid antibody fragments that recognize fluorescent proteins. Plants expressing these chimeric proteins were able to mount an immune response when fluorescent proteins were introduced. The work provides a way forward to rapidly and specifically modify susceptible crop species to provide them with pathogen resistance. —MRS

Science, abn4116, this issue p. 934

ACTIVE MATTER

Getting in your lane

Relatively sparse active systems, ranging in scale from atoms to pedestrian traffic, tend to form lanes when forced to cross from two different directions. Bacik *et al*. derived a new formulation for the angles of those lanes based on simple rules in the case of two intersecting groups of pedestrians. In an experimental arena, directing people to walk toward each other with simple rules, such as moving to the right to avoid a collision, generated a number of interesting laning scenarios. Backed by numerical simulations, the authors developed a general model for determining when laning occurs and how it will look depending on the microscopic variables of a system. —BG

Science, add8091, this issue p. 923

Models help explain how lanes emerge in active systems such as pedestrians walking in opposing directions, as in this busy intersection in Hong Kong.

MUSCULAR DYSTROPHY

Activating adenylyl cyclase to treat DMD

Duchenne muscular dystrophy (DMD) is characterized by repeated contraction-induced muscle injury that eventually leads to failure of skeletal muscle regeneration. Extraocular muscles (EOMs) are generally spared in DMD, suggesting the presence of a potential compensatory mechanism. Taglietti *et al.* generated a rat model of DMD and found that EOMs had increased thyroid stimulating hormone receptor (TSHR), resulting in decreased muscle stem cell senescence. Forskolin, an adenylyl cyclase activator, stimulated TSHR signaling and led to improved proliferation and reduced senescence of skeletal muscle stem cells, resulting in improved functional performance in affected rats. These findings suggest that activation of TSHR signaling may have therapeutic value in patients with DMD. —MN

Sci. Transl. Med. **15**, eadd5275 (2023).

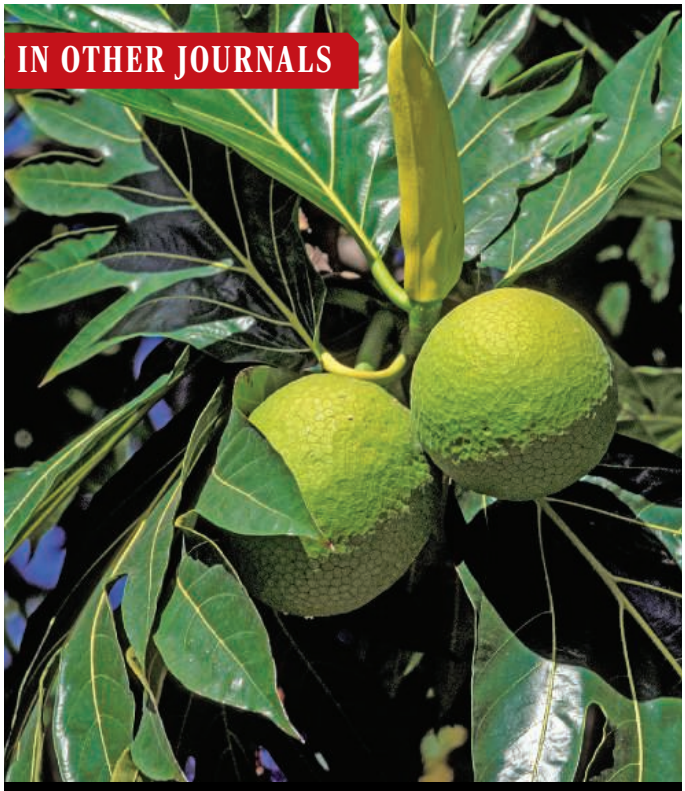
HEMATOPOIESIS

Trimming microRNA for blood development

Mutations in the 3′ to 5′ RNA exonuclease USB1 cause a pediatric disease with defects in the production of blood cells, although the underlying cause of this syndrome is unknown. Jeong *et al*. determined that USB1 removes extra adenosines from the 3’ end of microRNAs, which if not removed, promote degradation of microRNAs that are necessary for blood development. Blocking the enzyme that adds extra adenosines to microRNAs also restores the production of blood cells in mutant settings. This work determines the molecular basis of disease, identifies new roles for USB1, and suggests a possible therapeutic intervention for patients. —DJ

Science, abj8379, this issue p. 901

IN OTHER JOURNALS



PLANT BIOGEOGRAPHY

The terrible history of breadfruit

In 1789, Captain William Bligh’s attempt to export breadfruit plants from Tahiti to the Caribbean was foiled by mutiny. By 1793, breadfruit was successfully introduced to St. Vincent, from where it was distributed across the Caribbean as cheap food for enslaved people. Many island communities exposed to climate change now rely on breadfruit as a staple, so it is important to understand its genetics to enable appropriate management. Audi *et al.* combined genetic sequence analysis of 238 individual breadfruit with anatomical characterization, historical records, and local knowledge. The authors discerned five major Caribbean lineages of breadfruit, which match voyage records for the five types originally introduced from East Polynesia. Two Caribbean cultivars from St. Vincent and the Grenadines were confidently found to match extant Tahitian and Timorese cultivars genetically, but several other possible matches were identified. —MRS *Curr. Biol.* **33**, 287 (2023).

Breadfruit is a starchy food crop originally brought to the Caribbean to cheaply feed enslaved people.

ANCIENT DNA

Unearthing history in Charleston

Ancient DNA can be a powerful tool in uncovering the population history of individuals whose origins have been lost or intentionally obscured, as in the case of enslaved peoples. Working with community stakeholders, Fleskes *et al.* examined genomic and mitochondrial DNA from 36 individuals from the 18th century whose interment was discovered during construction in South Carolina. No known historical record of

PHOTOS: (LEFT TO RIGHT) WOLFGANG KAEHLER/LIGHTROCKET VIA GETTY IMAGES

EXTINCTION

Lost reptiles' role

The end of the Pleistocene saw a mass extinction among large mammals, which was attributed to the arrival of humans. Reptiles also play important roles in ecosystems in pollination and seed dispersal. Much less is known about how reptile extinctions may have affected ecosystems. Kemp has looked at changes in functional entities (FEs), which represent groups of reptiles with similar traits that provide similar ecosystem services across the Caribbean islands. As humans began to affect the archipelago during the Quaternary, the author found that more FEs were lost from smaller islands than larger ones, and that in some cases, species services contributed by FEs have been completely lost. Although more recent introductions have restored some ecological functions, in most cases these introduced species have left native FEs vulnerable to further loss. —SNV

Proc. Natl. Acad. Sci. U.S.A. **120**, e2201944119 (2023).

Green iguanas have been introduced in many Caribbean islands, providing some, but not all, functions of species lost due to human activities.

these individuals, known as the Anson Street Ancestors, exists, although evidence indicates that they were enslaved. Genetic analyses revealed that they were mostly of West and West-Central African descent, with one showing signatures of North American Indigenous ancestry and another being related to the Fulani populations of the Sahel region of Africa. —CNS

Proc. Natl. Acad. Sci. U.S.A. **120**, e2201620120 (2023).

ECOLOGY

Strix monitors for mammals

Tawny owls (*Strix aluco*) are common generalist predators found across Europe. This species maintains the same territories and the same mate year round. Yatsiuk *et al.* took advantage of a characteristic of tawny owls, which is to regurgitate pellets of indigestible prey remains. Examination of owl pellets deposited below roosting sites is a valuable way of surveying for small mammals that are hard to trap. The authors applied this method over 13 years around Kharkiv, Ukraine, to monitor small mammal activity through the seasons. A total of 23 species, ranging from moles to bats, were identified from pelleted fur and bones. Because the composition of the owl pellets reflects mammal species abundance within each territory, the data can reveal aspects of the more cryptic ecology of this forested steppe region. —CA

Biodivers. Data J. **11**, e98772 (2023).

MACHINE LEARNING

ML-assisted design of drug formulations

Polymer-based long-acting injectables (PLAIs), which release their therapeutic cargos over a prolonged period, are among the most promising drug delivery systems for the treatment of chronic diseases. However, because of multiple physicochemical factors that affect drug-polymer compatibility and cooperative performance, traditional approaches based on iterative trial and error become a significant obstacle in the development of their formulation. Bannigan *et al.* report that a machine learning (ML) approach trained with results from a light gradient boosting machine accurately predicted fractional drug release from various PLAI systems. This approach could be used to accelerate the design of new PLAIs, as demonstrated for a drug currently used to treat ovarian cancer. This work is an important step for ML approaches within pharmaceutical sciences and highlights the need for more open-source datasets to train such models. —YS

Nat. Commun. **14**, 35 (2023).

CRIMINAL JUSTICE

Leniency can improve public safety

Decisions to not prosecute defendants for nonviolent misdemeanor offenses reduced the likelihood of a subsequent criminal complaint against them by 53% compared with those who were prosecuted. Agan *et al.* studied 67,000 cases from 2000 to 2020 from Suffolk County, Massachusetts, where assistant district attorneys who vary in the leniency of their prosecution decisions are assigned these cases randomly. Defendants who were not prosecuted had 60% fewer criminal complaints against them over the next 2 years. Defendants without prior criminal records showed the largest effects, suggesting the important influence of having a criminal record on subsequent behavior. —BW

Q. J. Econ. 10.1093/qje/qjad005 (2023).

RADIATIVE COOLING

Freezing fresh water

Climate change is creating both water scarcity issues and the need for less energy-intensive strategies for freshwater production. Huang *et al.* demonstrate how passive radiative cooling can freeze and desalinate salt water. Their strategy uses outer space as a cold sink to help freeze the water, which then can be separated from the salty brine. This should work for any salinity level, does not require energy-intensive methods to produce fresh water, and can be combined with evaporative methods for 24-hour-a-day production. —BG

Joule **6**, 2762 (2022).

PHOTO: NHPA/PHOTOSHOT/SCIENCE SOURCE

RESEARCH ARTICLE

PLANT GENETICS

Dual domestications and origin of traits in grapevine evolution

Yang Dong[1,2]†, Shengchang Duan[1,2]†, Qiuju Xia[3]†, Zhenchang Liang[4]†, Xiao Dong[1,2]†§, Kristine Margaryan[5,6]‡, Mirza Musayev[7]‡, Svitlana Goryslavets[8]‡, Goran Zdunić[9]‡, Pierre-François Bert[10]‡, Thierry Lacombe[11]‡, Erika Maul[12]‡, Peter Nick[13]‡, Kakha Bitskinashvili[14]‡, György Dénes Bisztray[15]‡, Elyashiv Drori[16,17]‡, Gabriella De Lorenzis[18]‡, Jorge Cunha[19,20]‡, Carmen Florentina Popescu[21]‡, Rosa Arroyo-Garcia[22]‡, Claire Arnold[23]‡, Ali Ergül[24]‡, Yifan Zhu[1]‡, Chao Ma[25]‡, Shufen Wang[1,2], Siqi Liu[1,2], Liu Tang[1,2], Chunping Wang[1,2], Dawei Li[1,2], Yunbing Pan[1,2], Jingxian Li[1,2], Ling Yang[1,2], Xuzhen Li[1,2], Guisheng Xiang[1,2], Zijiang Yang[1,2], Baozheng Chen[1,2], Zhanwu Dai[4], Yi Wang[4], Arsen Arakelyan[5,26,27], Varis Kuliyev[28], Gennady Spotar[8], Nabil Girollet[10], Serge Delrot[10], Nathalie Ollat[10], Patrice This[11], Cécile Marchal[29], Gautier Sarah[11], Valérie Laucou[11], Roberto Bacilieri[11], Franco Röckel[12], Pingyin Guan[13], Andreas Jung[30], Michael Riemann[13], Levan Ujmajuridze[14], Tekle Zakalashvili[14], David Maghradze[14], Maria Höhn[15], Gizella Jahnke[15], Erzsébet Kiss[15], Tamás Deák[15], Oshrit Rahimi[16], Sariel Hübner[31], Fabrizio Grassi[32,33], Francesco Mercati[34], Francesco Sunseri[35], José Eiras-Dias[19,20], Anamaria Mirabela Dumitru[21], David Carrasco[22], Alberto Rodriguez-Izquierdo[22], Gregorio Muñoz[36], Tamer Uysal[37], Cengiz Özer[37], Kemal Kazan[38], Meilong Xu[39], Yunyue Wang[1], Shusheng Zhu[1], Jiang Lu[40], Maoxiang Zhao[25], Lei Wang[25], Songtao Jiu[25], Ying Zhang[41], Lei Sun[41], Huanming Yang[42], Ehud Weiss[43], Shiping Wang[25], Youyong Zhu[1], Shaohua Li[4]*, Jun Sheng[1,2]*, Wei Chen[1,2]*

We elucidate grapevine evolution and domestication histories with 3525 cultivated and wild accessions worldwide. In the Pleistocene, harsh climate drove the separation of wild grape ecotypes caused by continuous habitat fragmentation. Then, domestication occurred concurrently about 11,000 years ago in Western Asia and the Caucasus to yield table and wine grapevines. The Western Asia domesticates dispersed into Europe with early farmers, introgressed with ancient wild western ecotypes, and subsequently diversified along human migration trails into muscat and unique western wine grape ancestries by the late Neolithic. Analyses of domestication traits also reveal new insights into selection for berry palatability, hermaphroditism, muscat flavor, and berry skin color. These data demonstrate the role of the grapevines in the early inception of agriculture across Eurasia.

The cultivated grapevine (*Vitis vinifera* ssp. *vinifera*, hereafter *V. vinifera*) shares a close relationship with humans (*1*). With unmatched cultivar diversity, this food source (table and raisin grapes) and winemaking ingredient (wine grapes) became an emblem of cultural identity in major Eurasian civilizations (*1*–*3*), leading to intensive research in ampelography, archaeobotany, and historical records to reveal its history (*4*). Early work asserted that *V. vinifera* originated from its wild progenitor *Vitis vinifera* ssp. *sylvestris* (hereafter *V. sylvestris*) ~8000 years ago during the Neolithic agricultural revolution in the Western Asia (*5*, *6*). In recent years, various genetic studies explored this proposition (*6*–*13*), but the critical details of grapevine domestication were often inconsistent. Studies argued for the existence of domestication centers in the western Mediterranean (*13*), Caucasus (*12*, *14*), and Central Asia (*12*), which in turn cast doubt on the popular notion of a single past domestication event (*10*, *11*). Three demographic inferences yielded population split times between *V. vinifera* and *V. sylvestris* to dates between 15,000 and 400,000 years ago, predating the historical consensus on domestication time (*7*–*9*). Because early domesticates spread to other parts of Eurasia through poorly defined migration routes in the ensuing millennia (*5*), the single-origin theory also confounds the origin order between table and wine grapevines. One view proposes a wine grapevine–first model, with the two types diverging ~2500 years ago (*7*, *10*, *11*). Hybridization with local *V. sylvestris* was common in creating extant European wine grapes (*10*, *11*), but when these introgression events occurred is unknown. Several studies suggest that the earliest cultivation of European wine grapes in France and Iberia postdates 3000 years ago (*10*, *15*). These discrepancies primarily result from the inadequate sampling of grapevine accessions and the limited resolution of genetic data in previous analyses. Therefore, we report the genomic variation dataset from a global cohort to systematically delineate the structure of grapevine genetic diversity, explore the origin of *V. vinifera*, deduce a putative dispersal history, and investigate key domestication traits and diversification signatures.

Results

We constructed a chromosome-level reference *V. sylvestris* genome assembly (VS-1 from Tunisia) to attain genomic variations, which shows a higher percentage of anchored chromosomal lengths than PN40024 (fig. S1 and tables S1 to S9) (*16*). From the 3304 assembled accessions from a dozen Eurasian germplasm and private collections, we obtained good-quality Illumina paired-end sequencing data to an average 20× coverage for 3186 grapevine accessions (2237 *V. vinifera* and 949 *V. sylvestris*; tables S10 to S13). The sample selection preferentially included old, autochthonous, and economically important varieties to maximize the spectrum of genetic diversity. We also included genomic data for 339 previously sequenced accessions (266 *V. vinifera* and 73 *V. sylvestris*; table S14) in the analyses (*7*, *8*, *17*), producing the final cohort of 3525 grapevine accessions (2503 *V. vinifera* and 1022 *V. sylvestris*). The alignment of the Illumina reads to the VS-1 reference genome identifies 45,624,306 biallelic single-nucleotide polymorphisms (SNPs) and 7,314,397 biallelic short Indels [≤40 base pairs (bp); 73.2% shorter than 5 bp] (*16*), among which rare alleles (minor allele frequency ≤1%) accounted for the majority (fig. S2 and tables S15 to S22).

Core accessions differentiate by eight distinct genetic ancestries

Clones, mutants, synonyms, and homonyms are common phenomena in grapevine germplasm and collections (*18*). Using the identity-by-state sharing pattern estimators, we found 1534 accessions sharing the genetic profile with at least one other in the cohort, totaling 498 distinct genotypes (fig. S3 and table S23) (*16*). We kept one accession for each distinct genotype, corrected misidentified accessions, and excluded interspecific hybrids for a core cohort of 2448 grapevines (1604 *V. vinifera* and 844 *V. sylvestris*; fig. S3), which remain representative of the major viticultural regions (*19*) in the world (Fig. 1A and fig. S3).

Principal component analysis (PCA) showed that *V. sylvestris* and *V. vinifera* separately spread out along the first two axes (total variance explained: PC1 7.56% and PC2 1.71%), with both displaying a crude Western Asia to Western Europe gradient (Fig. 1B and figs. S4 and S5). The PC3 axis (1.26% variance) separates *V. vinifera* individuals according to their utilization, agreeing with the main table and wine grapevine clades in the maximum likelihood phylogenetic tree and reticulate phylogenetic network (figs. S6 and S7). The *V. vinifera* accessions show a weak isolation-by-distance correlation (Fig. 1C), suggesting a disconnection between the viticultural geographic pattern and the genetic structures in the grapevine

(*20*). This observation could be due to the extensive exchange of superior cultivars across regions and the subsequent interbreeding throughout history.

Given the poor resolution of viticultural regions in defining grapevine diversity, we leveraged genetic ancestry information from an unsupervised ADMIXTURE analysis to categorize core accessions (Fig. 1D and fig. S8) (*16*). At K = 2, all *V. vinifera* accessions contain a majority east (red) ancestry that matches the ancestry of the *V. sylvestris* accessions in the East Mediterranean region. At K = 8, hierarchical clustering of ancestry components identifies four *V. sylvestris* groups from distinct geographic regions: Western Asia (Syl-E1, 84.3% *K*2), the Caucasus (Syl-E2, 72.7% *K*6), Central Europe (Syl-W1, 94.7% *K*1), and the Iberian Peninsula (Syl-W2, 69.8% *K*8; Fig. 1, D to F). *V. sylvestris* accessions collected from other regions show admixed genetic structures (*16*). For cultivated grapevines (CGs), six genetic ancestries could designate six distinctive groups (CG1 to CG6), all covering a broad range of viticultural regions (Fig. 1, D to F) (*16*). Accessions with pure or close to pure ancestries (fig. S9) (*16*) helped to ascribe names to these groups as Western Asian table grapevines (CG1, 73.9% *K*2), Caucasian wine grapevines (CG2, 66.4% *K*6), muscat grapevines (CG3, 87.7% *K*5), Balkan wine grapevines (CG4, 69.9% *K*4), Iberian wine grapevines (CG5, 68.8% *K*7), and Western European wine grapevines (CG6, 68.4% *K*3). The admixed *V. vinifera* accessions showed different combinations of genetic ancestries (fig. S9). The four *V. sylvestris* and six *V. vinifera* groups, supported by archetypal analysis at K = 8 (fig. S10), formed identifiable clusters in the PCA plots (Fig. 1G and fig. S4) and were thus suitable for population genomic investigations.

Separation of V. sylvestris *ecotypes in Pleistocene*

According to the genetic ancestries and the occupied ecological niches in the western Eurasia continent, we designate *V. sylvestris* accessions in Western Asia and the Caucasus as the eastern ecotype (*V. sylvestris* eastern ecotype, hereafter Syl-E) and accessions in Central Europe and the Iberian Peninsula as the western ecotype (*V. sylvestris* western ecotype, hereafter Syl-W) (Fig. 2A). The large between-ecotype fixation index values [e.g., Syl-E1 versus Syl-W1, pairwise population fixation index (F_{ST}) = 0.340] and the small within-ecotype fixation index values (Syl-E1 versus Syl-E2, F_{ST} = 0.101; Syl-W1 versus Syl-W2, F_{ST} = 0.072; fig. S11 and table S26) support this designation. Both nucleotide diversity (π) and individual heterozygosity show that the western ecotype (especially Syl-W1) has significantly reduced variation compared with its eastern counterpart (fig. S11). Furthermore, the linkage disequilibrium decay (LD, r^2) was much slower in Syl-W (1.0 to 1.6 Kb at half of maximum r^2) than in Syl-E (400 to 600 bp at half of maximum r^2; fig. S12). These data demonstrate that the eastern ecotype retains more genetic diversity.

Demographic inference with folded SNP frequency spectra reveals an ancient population bottleneck in Syl-E ~400,000 to 800,000 years ago and in Syl-W ~150,000 to 400,000 years ago (Fig. 2B and fig. S13). This Pleistocene period, characterized by changing climate cycles (*21*, *22*), also witnessed the deduced population split (median time ~200,000 to 400,000 years ago) between the two ecotypes (Fig. 2C). The slow descent of the split line suggests that the geographic isolation process was gradual (fig. S13). At ~56,000 years ago, the population split between Syl-E1 and Syl-E2 occurred during the last glacial cycle (11,700 to 115,000 years ago), when the global climate trended toward dryer and colder conditions (*23*). Close to the time of the Last Glacial Maximum (LGM; ~21,000 years ago), *V. sylvestris* subgroups experienced a second population bottleneck (~40,000 years ago), with effective population sizes (N_e) reaching a minimum of 10,000 to 40,000 (Fig. 2B and fig. S13). After this result, ecological niche modeling predicts that the areas with suitable environmental conditions for Syl-E and Syl-W (suitability > 0.75) remained connected at the Pleistocene Last Interglacial (~130,000 years ago) (fig. S14) but became entirely separated at the LGM (Fig. 2D). The post-bottleneck N_e rebound was steeper in the Syl-W accessions, but the numbers decreased to lower levels in recent times (Fig. 2B and fig. S13). This result agrees with the reduced genetic diversity in Syl-W and the abrupt population split between Syl-W1 and Syl-W2 at ~2500 years ago.

Dual origin of V. vinifera *at the advent of agriculture*

The wet climate in the Early Holocene (11,700 to 8300 years ago) (*24*) facilitated the expansion of suitable habitats for Syl-E, resulting in a large geographic span from Central Asia to the Iberian Peninsula (Fig. 2D). This expansion supports the eastern origin and subsequent continental dispersal of *V. vinifera*. Because CG1 shares the main ancestral component with Syl-E1 and CG2 with Syl-E2 (Fig. 1, D and F), the possibility of two domestication events becomes evident. Indeed, both CG1 and CG2 maintain the highest genetic diversity and manifest the quickest LD decay among all CG groups (figs. S11 and S12). Furthermore, they are less differentiated from their corresponding wild ecotypes (Fig. 3A and fig. S11). The Akaike information criterion (AIC)–based phylogenetic selection also prefers a dual origin tree model (fig. S15), which agrees with the outgroup f_3 statistics biplots that CG1 and CG2 are genetically

[1]State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan Agricultural University, Kunming 650201, China. [2]Yunnan Research Institute for Local Plateau Agriculture and Industry, Kunming 650201, China. [3]State Key Laboratory of Agricultural Genomics, BGI-Shenzhen, Shenzhen 518083, China. [4]Beijing Key Laboratory of Grape Science and Oenology and Key Laboratory of Plant Resources, Institute of Botany, the Chinese Academy of Sciences, Beijing 100093, China. [5]Institute of Molecular Biology, NAS RA, 0014 Yerevan, Armenia. [6]Yerevan State University, 0014 Yerevan, Armenia. [7]Genetic Resources Institute, Azerbaijan National Academy of Sciences, AZ1106 Baku, Azerbaijan. [8]National Institute of Viticulture and Winemaking Magarach, Yalta 298600, Crimea. [9]Institute for Adriatic Crops and Karst Reclamation, 21000 Split, Croatia. [10]Bordeaux University, Bordeaux Sciences Agro, INRAE, UMR EGFV, ISVV, 33882 Villenave d'Ornon, France. [11]AGAP Institut, University of Montpellier, CIRAD, INRAE, Institut Agro Montpellier, 34398 Montpellier, France. [12]Julius Kühn Institute (JKI) – Federal Research Center for Cultivated Plants, Institute for Grapevine Breeding Geilweilerhof, 76833 Siebeldingen, Germany. [13]Botanical Institute, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany. [14]LEPL Scientific Research Center of Agriculture, 0159 Tbilisi, Georgia. [15]Hungarian University of Agriculture and Life Sciences (MATE), 1118 Budapest, Hungary. [16]Department of Chemical Engineering, Ariel University, 40700 Ariel, Israel. [17]Eastern Regional R&D Center, 40700 Ariel, Israel. [18]Department of Agricultural and Environmental Sciences, University of Milano, 20133 Milano, Italy. [19]Instituto Nacional de Investigação Agrária e Veterinária, I.P./INIAV-Dois Portos, 2565-191 Torres Vedras, Portugal. [20]Green-it Unit, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, 2780-157 Oeiras, Portugal. [21]National Research and Development Institute for Biotechnology in Horticulture, Stefanesti, 117715 Arges, Romania. [22]Center for Plant Biotechnology and Genomics, UPM-INIA/CSIC, Pozuelo de Alarcon, 28223 Madrid, Spain. [23]University of Lausanne, 1015 Lausanne, Switzerland. [24]Biotechnology Institute, Ankara University, 06135 Ankara, Turkey. [25]Department of Plant Science, School of Agriculture and Biology, Shanghai JiaoTong University, Shanghai 200240, China. [26]Armenian Bioinformatics Institute, 0014 Yerevan, Armenia. [27]Biomedicine and Pharmacy, RAU, 0051 Yerevan, Armenia. [28]Institute of Bioresources, Nakhchivan Branch of the Azerbaijan National Academy of Sciences, AZ7000 Nakhchivan, Azerbaijan. [29]Vassal-Montpellier Grapevine Biological Resources Center, INRAE, 34340 Marseillan-Plage, France. [30]Historische Rebsorten-Sammlung, Rebschule (K39), 67599 Gundheim, Germany. [31]Galilee Research Institute (Migal), Tel-Hai Academic College, 12210 Upper Galilee, Israel. [32]Department of Biotechnology and Biosciences, University of Milano-Bicocca, 20126 Milano, Italy. [33]NBFC, National Biodiversity Future Center, 90133 Palermo, Italy. [34]Institute of Biosciences and Bioresources, National Research Council, 90129 Palermo, Italy. [35]Department AGRARIA, University Mediterranea of Reggio Calabria, Reggio 89122 Calabria, Italy. [36]IMIDRA, Alcalá de Henares, 28805 Madrid, Spain. [37]Viticulture Research Institute, Ministry of Agriculture and Forestry, 59200 Tekirdağ, Turkey. [38]Queensland Alliance for Agriculture and Food Innovation, University of Queensland, St. Lucia, Queensland 4072, Australia. [39]Institute of Horticulture, Ningxia Academy of Agricultural and Forestry Sciences, Yinchuan 750002, China. [40]Center for Viticulture and Oenology, School of Agriculture and Biology, Shanghai JiaoTong University, Shanghai 200240, China. [41]Zhengzhou Fruit Research Institutes, CAAS, Zhengzhou 450009, China. [42]BGI-Shenzhen, Shenzhen 518083, China. [43]The Martin (Szusz) Department of Land of Israel Studies and Archaeology, Bar-Ilan University, 5290002 Ramat-Gan, Israel.

***Corresponding author. Email: wchenntr@gmail.com (W.C.); shengjun@dongyang-lab.org (J.S.); shhli@ibcas.ac.cn (S.L.)**

†These authors contributed equally to this work. ‡Institution contacts for biological samples. Ordered by country names. §Present address: Department of Chromosome Biology, Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany.

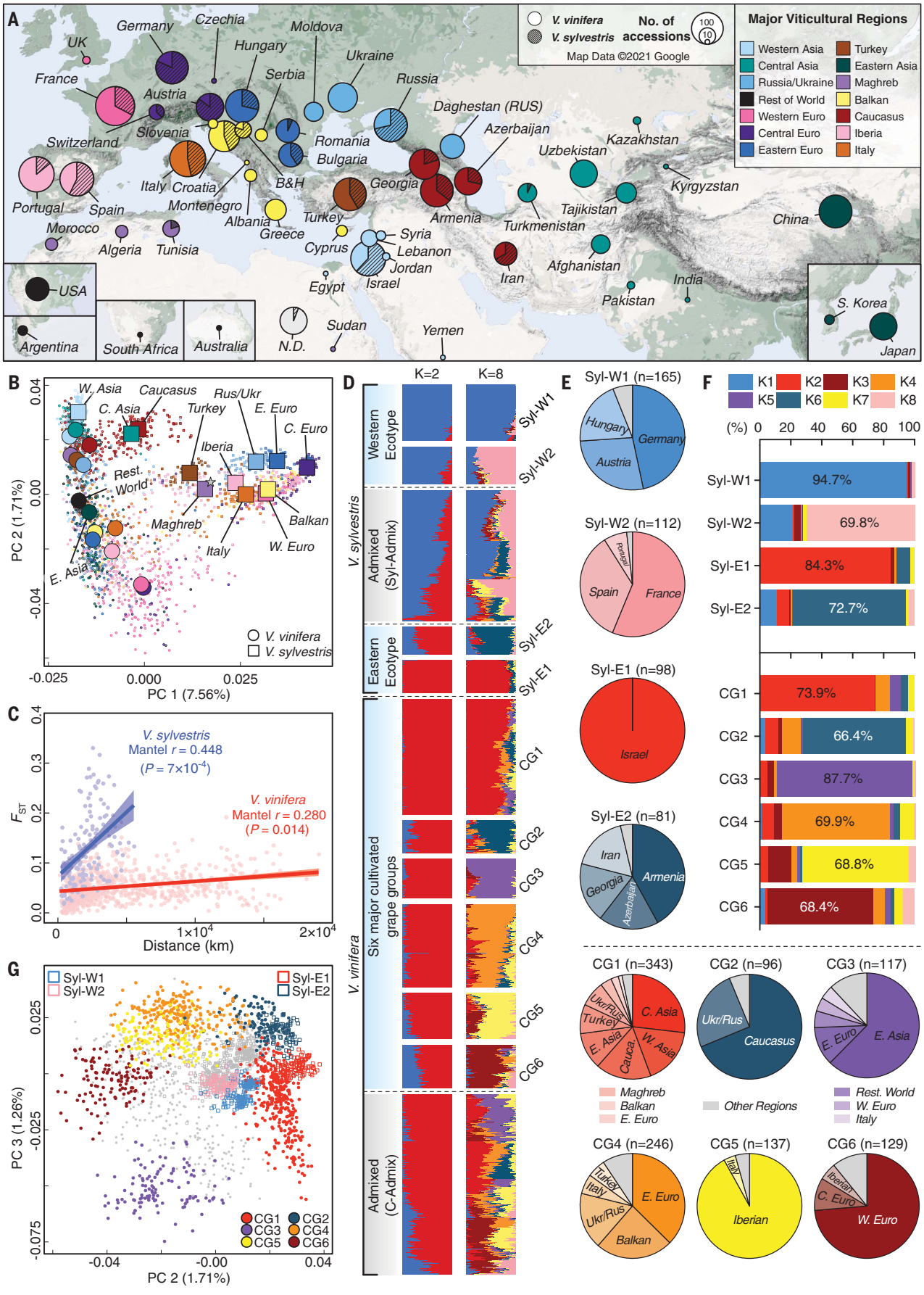

A
V. vinifera
V. sylvestris
No. of accessions
100
10
Map Data ©2021 Google
Major Viticultural Regions
Western Asia
Central Asia
Russia/Ukraine
Rest of World
Western Euro
Central Euro
Eastern Euro
Turkey
Eastern Asia
Maghreb
Balkan
Caucasus
Iberia
Italy
Germany
Czechia
Moldova
UK
Hungary
Ukraine
France
Serbia
Russia
Austria
Daghestan (RUS)
Azerbaijan
Switzerland
Slovenia
Romania
Kazakhstan
Bulgaria
Uzbekistan
Italy
Croatia
B&H
Georgia
Kyrgyzstan
Portugal
Spain
Montenegro
Albania
Turkey
Armenia
Tajikistan
China
Morocco
Greece
Turkmenistan
Algeria
Tunisia
Cyprus
Syria
Lebanon
Jordan
Afghanistan
Egypt
Israel
Iran
India
Pakistan
USA
S. Korea
Japan
Argentina
South Africa
Australia
N.D.
Sudan
Yemen
B
W. Asia
Caucasus
Rus/Ukr
C. Asia
Turkey
E. Euro
C. Euro
Iberia
Rest. World
Maghreb
Balkan
Italy
W. Euro
E. Asia
V. vinifera
V. sylvestris
PC 2 (1.71%)
PC 1 (7.56%)
C
V. sylvestris
Mantel r = 0.448
(P = 7×10-4)
V. vinifera
Mantel r = 0.280
(P = 0.014)
FST
Distance (km)
D
K=2
K=8
Western Ecotype
Admixed (Syl-Admix)
Eastern Ecotype
V. sylvestris
Six major cultivated grape groups
Admixed (C-Admix)
V. vinifera
Syl-W1
Syl-W2
Syl-E2
Syl-E1
CG1
CG2
CG3
CG4
CG5
CG6
E
Syl-W1 (n=165)
Hungary
Germany
Austria
Syl-W2 (n=112)
Portugal
Spain
France
Syl-E1 (n=98)
Israel
Syl-E2 (n=81)
Iran
Georgia
Armenia
Azerbaijan
F
K1
K2
K3
K4
K5
K6
K7
K8
(%) 0 20 40 60 80 100
Syl-W1
94.7%
Syl-W2
69.8%
Syl-E1
84.3%
Syl-E2
72.7%
CG1
73.9%
CG2
66.4%
CG3
87.7%
CG4
69.9%
CG5
68.8%
CG6
68.4%
CG1 (n=343)
Ukr/Rus
Turkey
C. Asia
E. Asia
W. Asia
Cauca.
CG2 (n=96)
Ukr/Rus
Caucasus
CG3 (n=117)
E. Euro
E. Asia
Maghreb
Balkan
E. Euro
Other Regions
Rest. World
W. Euro
Italy
CG4 (n=246)
Turkey
Italy
Ukr/Rus
E. Euro
Balkan
CG5 (n=137)
Italy
Iberian
CG6 (n=129)
Iberian
C. Euro
W. Euro
G
Syl-W1
Syl-W2
Syl-E1
Syl-E2
CG1
CG2
CG3
CG4
CG5
CG6
PC 3 (1.26%)
PC 2 (1.71%)

Fig. 1. Genetic diversity of global core *V. sylvestris* and *V. vinifera* accessions. (**A**) Geographical locations of the 2448 core grapevine accessions. (**B**) PCA according to major viticultural regions. Large square/circle highlights median position. Star shows VS-1 position. (**C**) Isolation-by-distance test of *V. sylvestris* and *V. vinifera* accessions. Linear regression with 95% confidence interval is shown. (**D**) ADMIXTURE clustering of the accessions. (**E**) Geographic locations of the accessions in each group. Gray represents minor locations. (**F**) Average proportion of major genetic ancestries in grapevine groups. (**G**) PC2 versus PC3 projection according to grapevine group.

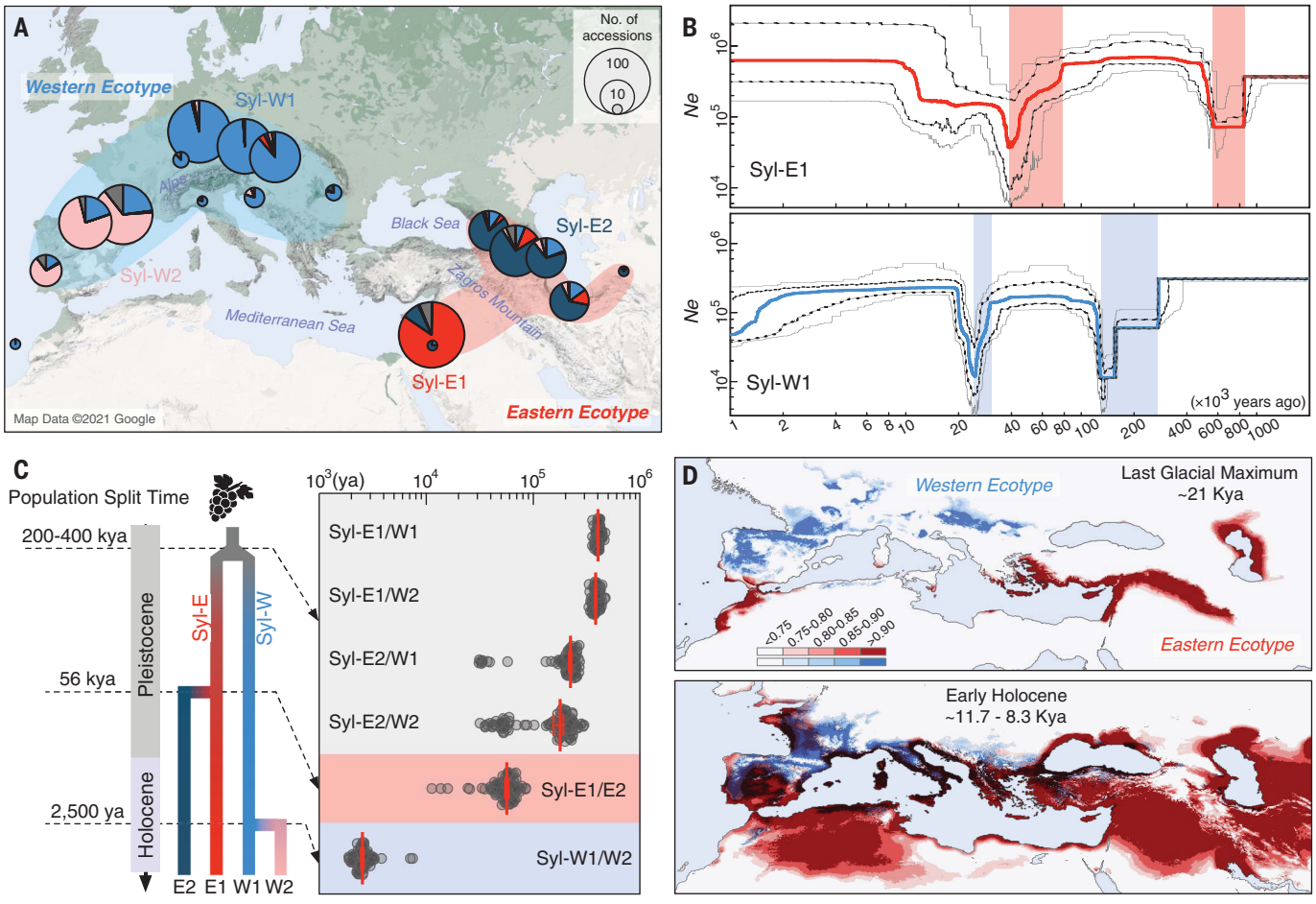


Fig. 2. Population history of *V. sylvestris* ecotypes. (**A**) Geographic isolation and population separation of *V. sylvestris* ecotypes. Pie charts show mean ancestry proportion at $K = 8$. Same color scheme as in Fig. 1B is used. (**B**) Demographic histories of *V. sylvestris* populations deduced from Stairway Plot 2. Lines indicate medians with 75% and 95% confidence intervals. (**C**) Population split times among ecotypes with MSMC2. Red bars indicate medians with 95% confidence intervals. (**D**) Ecological niche modeling of the suitable habitats for *V. sylvestris* ecotypes. The color scale shows suitability score.

closer to Syl-E1 and Syl-E2, respectively (Fig. 3B, fig. S15, and table S27). The population split lines of CG1/Syl-E2 and CG2/Syl-E1 pairs resemble that of Syl-E1/Syl-E2 and differ from those of CG1/Syl-E1 and CG2/Syl-E2 pairs (Fig. 3C and fig. S16). These data collectively support a dual origin of *V. vinifera* and reject the popular theory of a single primary domestication center (*10*, *11*). Both CG1/Syl-E1 and CG2/Syl-E2 population pairs separated quickly (Fig. 3C), which is compatible with a clean-split scenario. We estimate the median population split time to be ~11,000 years ago (95% confidence interval: ~10,500 to 12,500 years ago) for both pairs, suggesting that the domestication events took place concurrently around the advent of agriculture. Because CG1 and CG2 separately represent table and wine grapevine ancient genetic backgrounds (*K2* and *K6*; fig. S9), the dual origin rejects the assumption that wine grapevines predate table grapevines (*7*, *10*, *11*).

Dispersal of grapevine domesticates along human migration routes

The geographic distributions of CG1 and CG2 cultivars across Eurasia and North Africa correspond to vastly different human migration routes for the two grapevine groups (Fig. 3D). The CG2 cultivars were mainly confined to both sides of the Caucasus Mountains, with a limited dispersal into the Carpathian Basin by the northern Black Sea. This result contrasts with previous models implying that CG2 played a central role in the formation of wine grapevines in Europe (*3*). Instead, CG2 represents a local domestication effort that had a minor impact on grapevine diversification. By comparison, the dispersal of CG1 in four directions spanned Eurasia and North Africa. First, the eastward expansion through Central Asia into India and China follows the Inner Asian

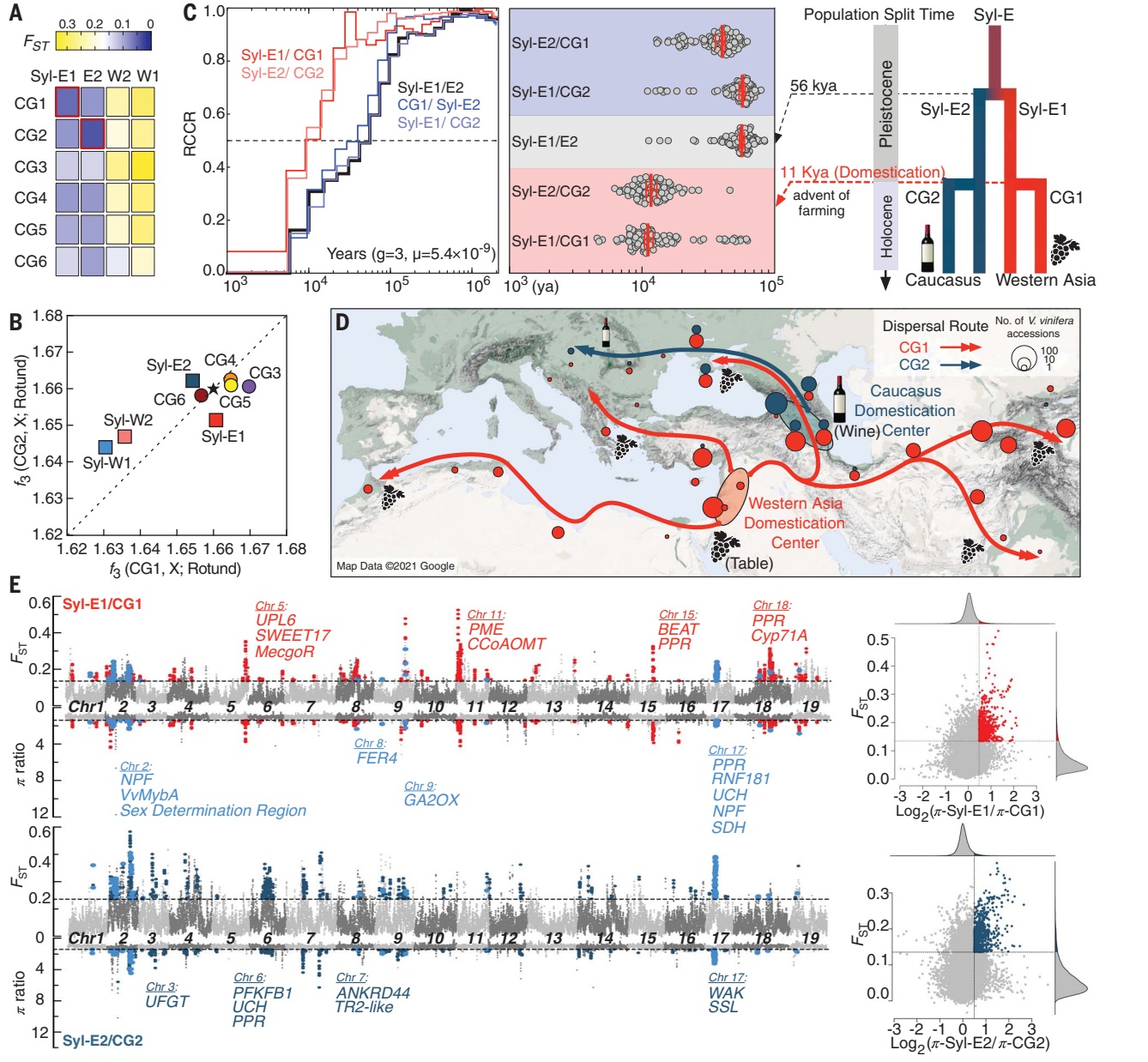


Fig. 3. Dual domestications of *V. vinifera* in Western Asia and the Caucasus. (**A**) Pairwise fixation index of the major grapevine groups. (**B**) Outgroup f_3 statistics biplot measuring genetic similarity. Rotund, *Muscadinia rotundifolia*. Stars mark the f_3 statistics for CG1/CG2. (**C**) Estimated split times among Syl-E1/2 and CG1/2 with MSMC2 (left). Red bars indicate medians with 95% confidence intervals. (**D**) Geographic distribution of CG1 and CG2 in relation to the domestication centers. Human dispersal routes are shown. (**E**) Shared (sky blue) and unique domestication selective sweep regions (red and dark teal) in *V. vinifera*.

Mountain Corridor, a path that also witnessed the exchange of other crops (i.e., wheat, barley, and millet) between the West and the East (*25*). Second, the northbound expansion could mirror the early cultural contact of Western Asia over the Zagros mountains with the Caucasus (*26*, *27*). Third, the northwest expansion through Anatolia into the Balkans bespeaks the spread of farming into Europe (*28*, *29*). Finally, a westward expansion moved across the North African coastline to reach Morocco (*30*). Even though grapevine domestication followed the trails of past human migration, the timing and dispersal details require paleogenomic data for delineation.

Shared and unique domestication signatures in CG1 and CG2 grapevines

Given the dual origin scenario, we investigated domestication signatures in both Syl-E1/CG1 and Syl-E2/CG2 group pairs by selecting genomic regions that display increased nucleotide diversity differences and population differentiation (both top 5%; Fig. 3D). This method yields 1140 domestication selective sweep genes in 132 regions for CG1 and 887 genes in 137 regions for CG2 (table S28), among which only 189 genes in 31 regions exist in both groups (table S29). Most shared signals are on chromosomes 2 and 17, confirming previous findings that the selection on flower sexual morphs (sex determination region, SDR), berry skin color (*VvMybA* gene cluster), and berry development (*SDH* gene cluster) were of great importance during grapevine domestication (*8*, *11*). In addition, our analysis identifies

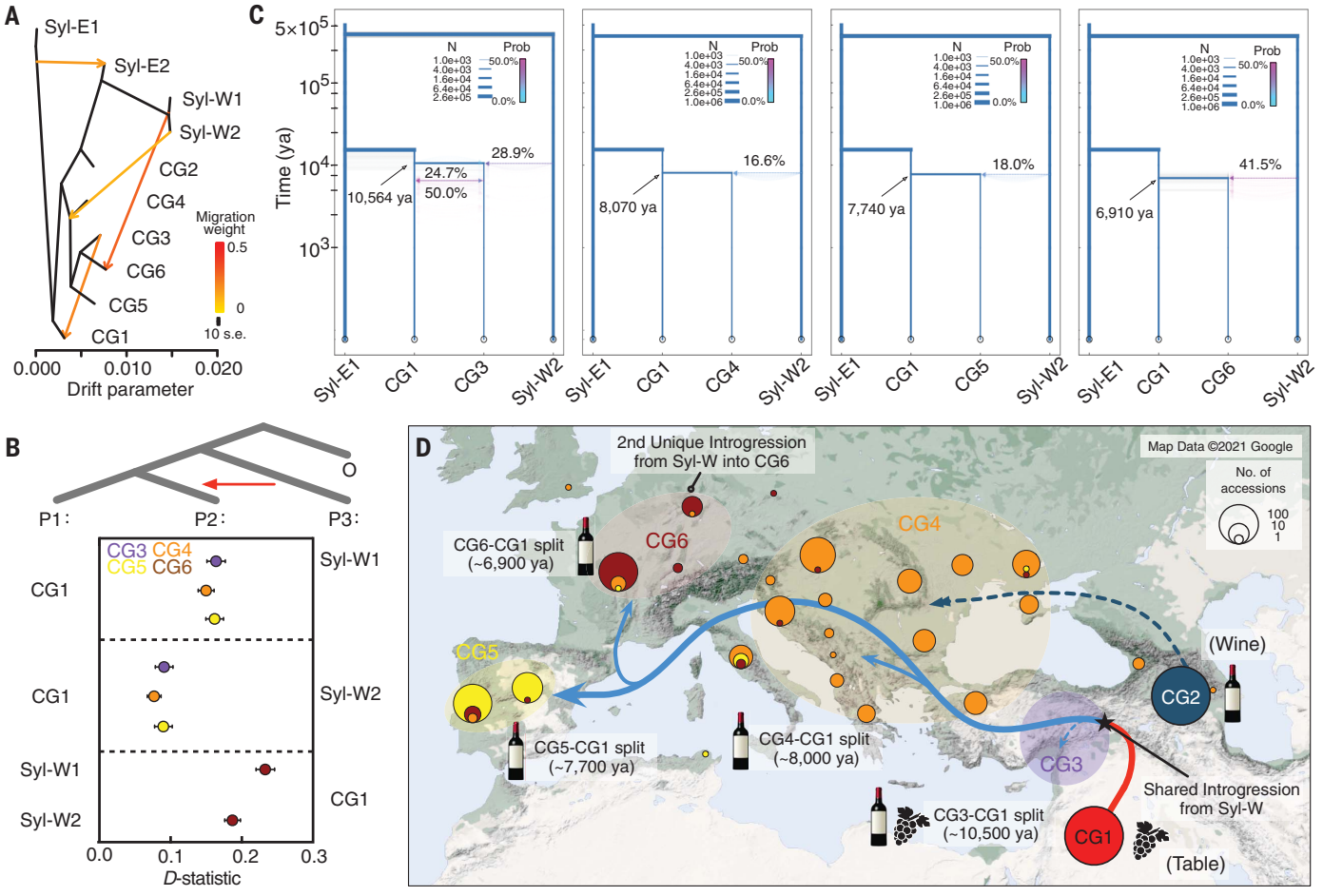


Fig. 4. Stepwise diversification of *V. vinifera* in Europe. (**A** and **B**) Introgression from Syl-W into European *V. vinifera* groups revealed by TreeMix (A) and confirmed by D-statistic (B). (**C**) Four population simulation of split times and genetic introgression using Momi2. Median numbers from 100 bootstrap runs are shown. (**D**) Origination of *V. vinifera* groups (CG3 to CG6) by the end of the Neolithic. Geographic distributions of CG groups are shown by colored circles. See fig. S24 for details on CG3.

shared domestication genes that possibly underlie grapevine growth (e.g., *NPF*), physiology (e.g., *FER4*), fruit set (e.g., the *GA2OX* gene cluster), and resistance to biotic/abiotic stress (e.g., *FER4*, the *PPR* gene cluster, and the *RNF181* gene cluster) [see (*16*) for gene descriptions].

As expected for dual domestications, most selective sweep signatures in CG1 and CG2 are unique and target distinctive chromosomal regions (Fig. 3E). Even though CG1 and CG2 correspondingly represent table and wine grapevines, many unique signatures seem to suggest a convergent selection mechanism targeting different aspects of common domestication traits. An obvious example is the improvement of berry palatability through the reduction of alkaloid biosynthesis (the *MecgoR* gene cluster in CG1 and the *TR2* and *SSL* gene clusters in CG2) and the enhancement of carbohydrate metabolism (*SWEET17* in CG1 and *PFKFB1* in CG2). Other examples include perceived berry desirability (the *BEAT* gene cluster for floral scent in CG1 and the *UFGT* gene cluster for berry color in CG2) and response to environmental stresses (*UPL6* in CG1 and *WAK* in CG2). These findings suggest that the initial cultivation of CG1 and CG2 may have been to serve early humans' caloric and micronutrient needs. The selection of genetic features suitable for winemaking in CG2 could have been serendipitous, and the practice of winemaking with CG2 (e.g., 8000 years ago) (*14*) possibly postdates grapevine domestication. Because gene annotation depends on homology-based inference, it should be noted that many genes mentioned here need further verification in grapevines.

Wine grapevine diversification in Europe

Because the CG1 early domesticates dispersed into Europe through Anatolia, a crucial question concerns the diversification history of European wine grapevines in the ensuing millennia. In particular, the shared areas of suitable habitats for Syl-E and Syl-W in the early Holocene (black area in Fig. 2D) formed an ecological foundation for the genetic exchange between CG1 and local refugia Syl-W accessions in the coastal regions of the northern Mediterranean Sea and the southern Black Sea, the Iberian Peninsula, and an area corresponding to present-day western France. It is therefore important to examine where and how distinct grapevine genetic ancestries (CG3 to CG6) formed with relevance to Syl-W introgression (*10*, *11*). We have chosen cultivars in each group with at least 75% major ancestry (and with an average Syl-W ancestry in each *V. vinifera* group <3%) to perform population analyses. This selection rules out many old varieties (i.e., 'Lambrusco' cultivars deriving about half of their ancestries from Syl-W; fig. S9), which likely showcase secondary diversification efforts after the distinct ancestries had been established. The TreeMix analysis finds one migration edge that points from Syl-W to a population ancestral to CG3 to CG6 (estimated weight, 0.114; Fig. 4A and fig. S17), suggesting an ancient introgression event occurred before the diversification of all European grapevines. An additional migration

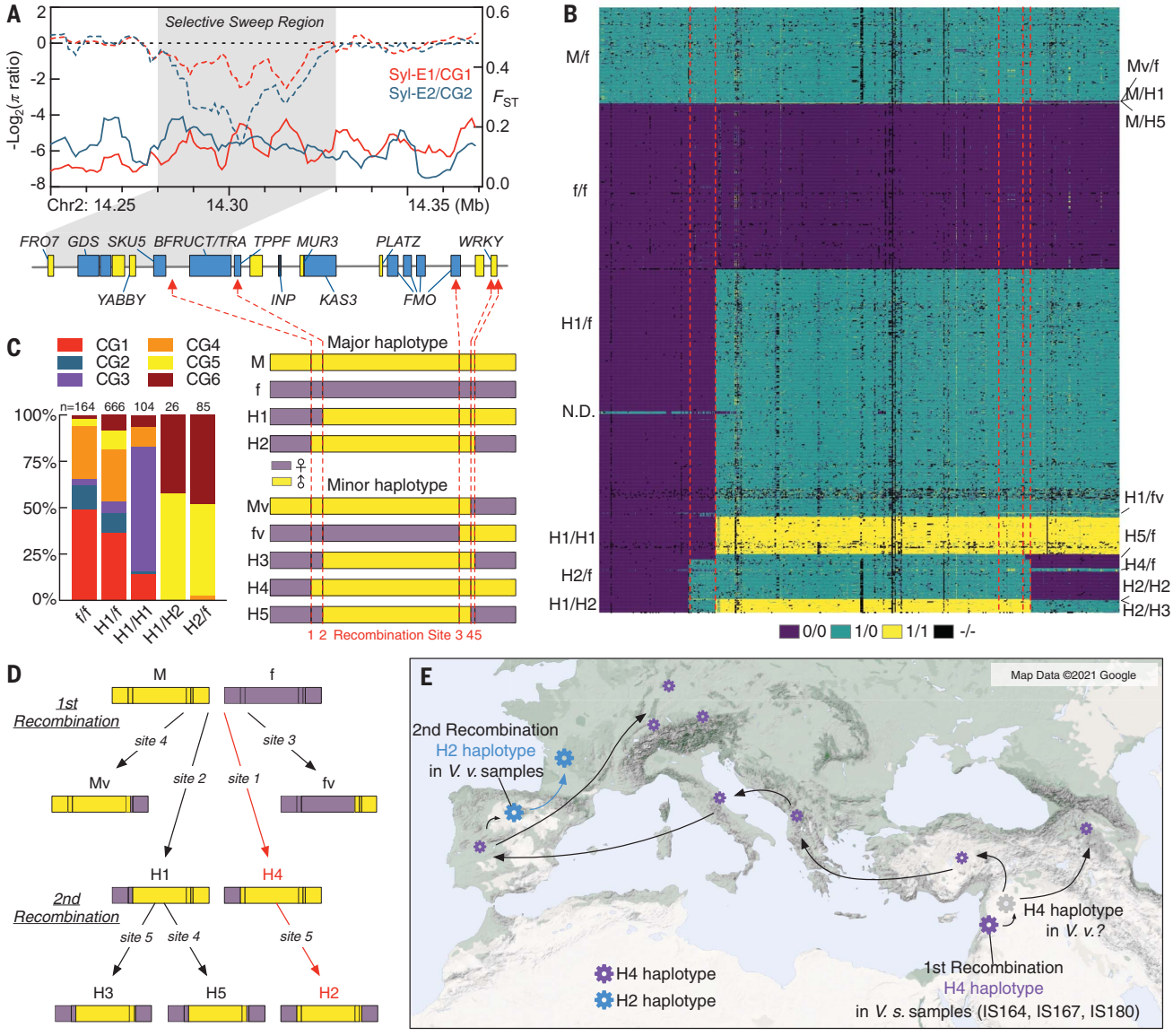


Fig. 5. Selection and evolution of the SDR in the core grapevine accessions. (**A**) The SDR in VS-1. Red arrows indicate identified recombination sites. (**B**) SDR genotypes from associated SNPs reveal five recombination sites (dashed lines) and genotype diversity (right). Major and minor haplotypes are shown on the left. (**C**) Distribution of SDR genotypes in the six major grapevine groups. (**D**) Recombination history of all SDR haplotypes. (**E**) Putative dispersal route of the H4 haplotype and the origination of H2 haplotype.

edge also points from Syl-W to CG6 (estimated weight, 0.292), which implies an independent introgression event unique to Western European wine grapevines in the past. Various combinations of D-statistics testing the gene flow from Syl-W into CG groups (Z score > 3.0, adjusted $P < 4.17 \times 10^{-5}$; Fig. 4B and table S31) support this introgression history. Additionally, gene flow from Syl-W into CG3 to CG6 inferred from Momi2 align with their corresponding divergence from CG1, further supporting the introgression history (Fig. 4C). The estimated median divergence times date the creation of muscat grapes (CG3) to 10,500 years ago, Balkan wine grapes (CG4) to 8070 years ago, Iberian wine grapes (CG5) to 7740 years ago, and Western European wine grapes to 6910 years ago (Fig. 4D). These stepwise diversification times agree with the historical migration of Anatolian farmers into Europe (*26*, *29*, *31*, *32*), substantiating the role of viticulture in forming Neolithic agricultural societies.

The migration edge weights, f_4 ratio, and Momi2 estimates collectively show that ancient introgression from Syl-W accounts for ~11.4 to 18.0% of the CG3 to CG6 genomes (Fig. 4 and table S30). In addition, at least one other independent introgression event contributed ~25.0 to 30.0% additional Syl-W to the CG6 ancestry. We have screened the introgression tracts in CG3 to CG6 by choosing the genomic windows with the top 1% d_f and f_{dM} values (fig. S18). Ten shared regions among the CG3 to CG6 groups contain genes that are putatively involved in plant immunity (e.g., *CYSK*), abiotic stress response (e.g., *GBA*), and carbohydrate metabolism (e.g., *TPS/TPP*) (table S31). This result agrees with the proposal that introgression helps grapevines adapt to new environments and become more suitable for winemaking (*10*, *11*).

Genetic analyses of domestication and diversification traits

Hermaphroditism: origin of H2 haplotype

The transition from dioecy in *V. sylvestris* (male, M/f; female, f/f) to hermaphroditism in *V. vinifera*

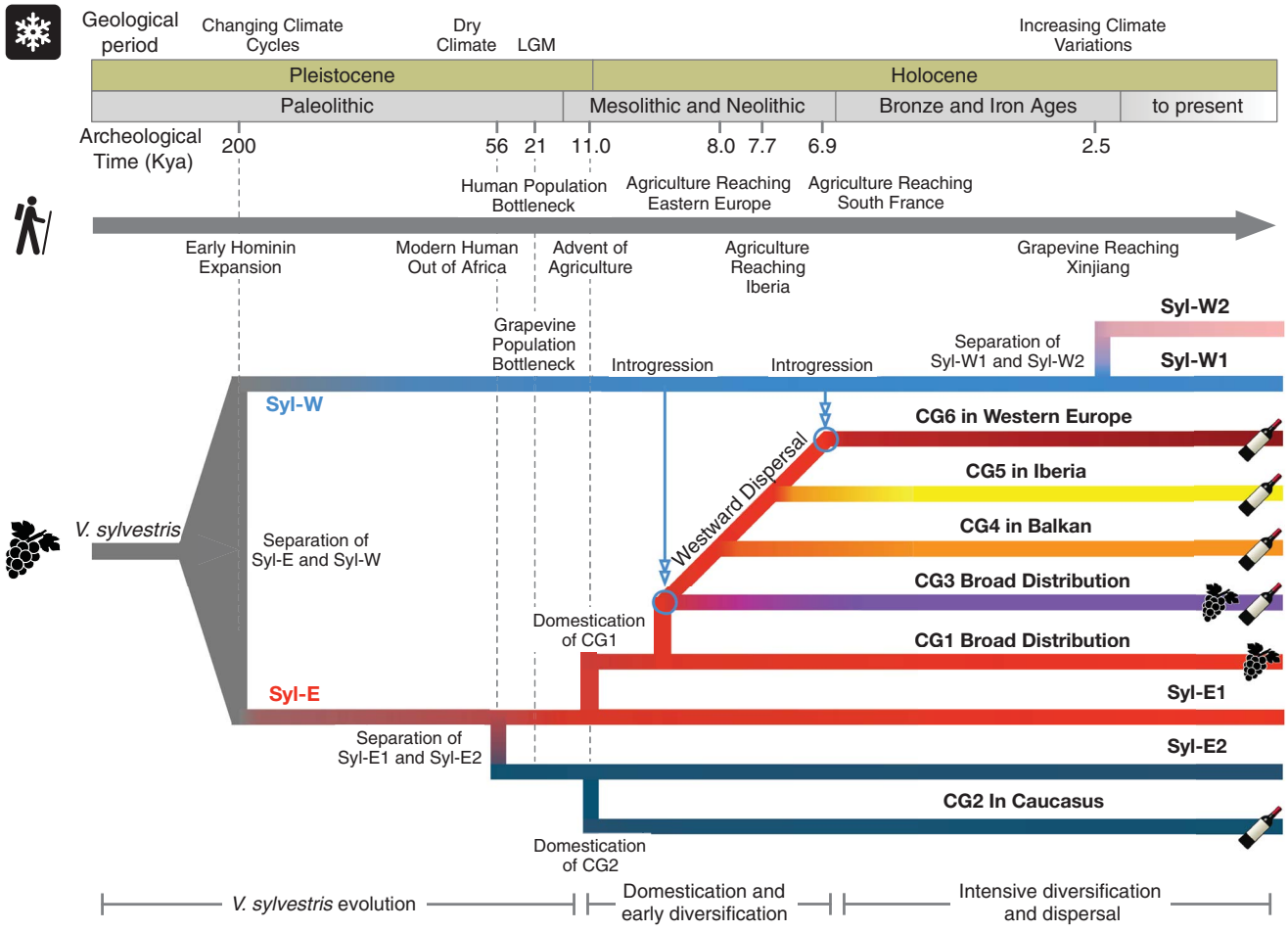


Fig. 6. Schematic graph of grapevine evolutionary history. Key events in the evolutionary history of grapevines are shown alongside major events in global climate change and human migration.

is the most prominent phenotypic change during domestication (*33*). It involves recombination events between M and f around a selective sweep region on chromosome 2 known as the SDR (Fig. 5A). Previous studies have identified two major hermaphroditic haplotypes (H1 and H2) and four hermaphroditic genotypes (H1/f, H2/f, H1/H1, and H1/H2) from select cultivars (*33*), but the recombination history remains unclear. The analysis of our grapevine cohort reveals five recombination sites in the SDR (Fig. 5B), which not only confirms known genotypes but also identifies new minor haplotypes (male variant Mv, female variant fv, H3, H4, and H5) and genotypes (Mv/f, M/H1, M/H5, H1/fv, H5/f, H4/f, H2/H2, and H2/H3) in both wild and cultivated grapevines (Fig. 5B and table S32). Among all SDR haplotypes, M and H1 manifest the highest subtype diversity (figs. S19 to S22). Furthermore, the SDR genotype statistics reveal a distribution bias of the H2-containing SDRs in the Iberian (CG5) and Western European (CG6) grapevines (Fig. 5C and fig. S23). To investigate this observation, we constructed a putative recombination history for all known SDR haplotypes (Fig. 5D), which showed that a first recombination event between the parental M and f haplotypes created Mv (site 4), fv (site 3), H1 (site 2), and H4 (site 1). On this basis, H1 experienced a second recombination event with f to produce H3 (site 5) and H5 (site 4), whereas H4 recombined again with f at site 5 to bring about H2. Because three Syl-E *V. sylvestris* (IS164, IS167, and IS180) and 11 *V. vinifera* accessions in the cohort contain H4 (Fig. 4G), a likely scenario supports a westward dispersal of H4 after human selection to reach the Iberian Peninsula [e.g., in extant old Iberian cultivar 'Malvasia Fina' (PO153)], where H2 originated from H4 through secondary recombination and later became dominant during the diversification of Iberian and Western European cultivars.

Muscat flavor: Trait selection may reduce grapevine fitness

Muscat grapevine is unique for its floral aromas, which result from a hard-to-define concoction of monoterpenoids in the fruit (*34*). Given the broad geographic distribution (fig. S24) and ancient history of muscat grapevines, it is not easy to pinpoint the center of origin. However, Momi2 estimate predicts a population split from CG1 at ~10,564 years ago (Fig. 4C), suggesting an origination site within the boundary of Western Asia. This scenario agrees with the relatively low F_{ST} values and sizeable gene flow with CG1 (Fig. 4 and fig. S11). The CG3 group also shows low genetic diversity and high LD extent compared with the others (figs. S11 and S12). One possible reason is the gradual loss of ancient CG3 cultivars in Anatolia and the surrounding regions throughout history (fig. S24). Even though the muscat aroma is a complex trait, genome-wide association analysis based on a binary differentiation reveals 18 SNP signatures on chromosomes 5 and 18 (fig. S24 and table S33). This set includes a nonsynonymous SNP Chr5:19419686 in the *VvDXS* gene linked to the trait (*34*). Examination of the genotype at this locus shows that 108 of the 134 muscat grapevines (including 'Muscat Hamburg,' 'Königin der Weingärten,' and 'Muscat of Alexandria,' which are commonly

used as parental cultivars) are heterozygous (G/T), and only eight individuals are homozygous (T/T) for the alternative SNP (exact test for Hardy-Weinberg equilibrium, D = 20.68, $P = 2.01 \times 10^{-13}$). Additionally, most grapevines without muscat aroma are homozygous for the reference SNP (G/G; 1451 of 1468; exact test for Hardy-Weinberg equilibrium, D = 0.049, $P = 1.00$). This result suggests that selection on this allele might have put a constraint on grapevine fecundity, thereby preventing the alternative SNP from reaching fixation.

Berry skin color: New genes associated with white grapes

The emergence of white grapes from their red-berried congeners is an essential domestication episode in viticulture history. The color change results from a reduction of anthocyanin synthesis in berry skin cells, where the expression of proposed master regulators such as *VvMybA* decreased significantly in select cultivars because of a *Gret1* retrotransposon (*35*), nonconservative exonic mutations (*36*), or large deletions in the locus (*37*). We performed genome-wide association analysis on this large grapevine cohort (fig. S25, A and B) and identified multiple significant SNPs across the genome (fig. S25C). The most prominent peak spans a broad genomic region from 3.51 to 16.05 Mb on chromosome 2, overlapping the *VvMybA* locus. Among all significant exonic SNPs in this region (table S34), nonsynonymous SNPs with the smallest *P* values localize to two uncharacterized genes outside the *VvMybA* locus (fig. S25D), the putative protein functions of which are acylaminoacyl-peptidase (*Vvsyl02G000229*) and lysine-specific demethylase (*Vvsyl02G001064*). These SNPs are overwhelmingly homozygous for the reference allele in white grapes and are heterozygous in red grapes (fig. S25E). We validated the SNPs in red-berried *V. sylvestris* accessions to account for possible false positives and confirmed their genotypes as being predominantly heterozygous (fig. S25E and table S34). By comparison, significant exonic SNPs in *VvMybA* genes [including Chr2:5116947 G/T reported previously in (*36*)] show shared genotypes between white grapes and the *V. sylvestris* accessions (fig. S25E). It is unclear how *Vvsyl02G000229* and *Vvsyl02G001064* might regulate anthocyanin synthesis, but these results demonstrate that exonic mutations in the two genes are better predictors of berry skin colors. Furthermore, the heterozygous SNP states in *V. sylvestris* accessions suggest that the white berry alleles existed in natural wild populations before grapevine domestication.

Discussion

Our systematic genomic survey of *V. sylvestris* and *V. vinifera* accessions paints a defined picture of grapevine evolutionary history, which echoes key events in the history of world climate change and human migration (Fig. 6). The Pleistocene era witnessed the continuous fragmentation of habitats, the decline of effective population size, and the separation of ecotypes for *V. sylvestris*. It is highly likely that modern humans extensively used grapevines as an energy source from the late Pleistocene, but the harsh climate was not suited for agriculture (*38*). As the climatic conditions ameliorated at the Pleistocene-Holocene transition, the grapevine, with its relatively stable perennial yield, unsurprisingly became one of the earliest candidates for domestication. The dual events underpin the model that plant domestication occurs in large, culturally connected areas over a long time (*39*), but the domestication time gap remains between genomic inference and archaeological evidence (table S35 and figs. S26 and S27) (*16*). The diverse SDR haplotypes suggest that an early goal could be the conscious selection (*40*) and propagation of rare, naturally occurring hermaphroditic individuals from the *V. sylvestris* population because they allow mass plantation without male plants. The selection on phenotype, but not on genotype, also implies that the different hermaphroditic haplotypes were subject to strong genetic drift, which is supported by the high frequency of H1 and the almost extinct H4 in extant cultivars. The Mesolithic and Neolithic periods also saw the early dispersal and diversification of grapevines such that unique ancestries emerged in the Balkans, Iberia, and Western Europe with the help of *V. sylvestris* introgression into CG1. This event mirrors early farmer migration in Europe, consolidating the role of viticulture in forming sedentary societies. A higher level of cultural exchange characterizes the last stage since the Bronze Age and the trading of superior grapevine cultivars along trade routes. This is especially evident in the plethora of Italian cultivars with three or more genetic ancestries, but unfortunately poses a challenge to disentangle the genealogical history of each grapevine cultivar (*20*). Finally, genetically reliable wild grapevines from Central Asia, a region battered by climate change and social instability for the past few millennia, are no longer available to test Vavilov's theory for a diversity center or a hypothetical turnover of grapevine types caused by Islam conversion in the region. Paleogenomic data may help to resolve these questions in the future.

REFERENCES AND NOTES

1. P. E. McGovern, U. Hartung, V. R. Badler, D. L. Glusker, L. J. Exner, *Expedition* **39**, 3–21 (1997).
2. P. This, T. Lacombe, M. R. Thomas, *Trends Genet.* **22**, 511–519 (2006).
3. F. Grassi, G. De Lorenzis, *Int. J. Mol. Sci.* **22**, 4518 (2021).
4. D. Cantu, M. A. Walker, *The Grape Genome* (Springer Nature, 2019).
5. D. Zohary, M. Hopf, E. Weiss, *Domestication of Plants in the Old World: The Origin and Spread of Domesticated Plants in Southwest Asia, Europe, and the Mediterranean Basin* (Oxford Univ. Press, 2012).
6. S. Myles *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3530–3535 (2011).
7. Y. Zhou, M. Massonnet, J. S. Sanjak, D. Cantu, B. S. Gaut, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11715–11720 (2017).
8. Z. Liang *et al.*, *Nat. Commun.* **10**, 1190 (2019).
9. A. Sivan *et al.*, *Plants People Planet* **3**, 414–427 (2021).
10. S. Freitas *et al.*, *Sci. Adv.* **7**, eabi8584 (2021).
11. G. Magris *et al.*, *Nat. Commun.* **12**, 7240 (2021).
12. S. Riaz *et al.*, *BMC Plant Biol.* **18**, 137 (2018).
13. R. Arroyo-García *et al.*, *Mol. Ecol.* **15**, 3707–3714 (2006).
14. P. McGovern *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E10309–E10318 (2017).
15. J. Ramos-Madrigal *et al.*, *Nat. Plants* **5**, 595–603 (2019).
16. See the supplementary materials.
17. M. J. Roach *et al.*, *PLOS Genet.* **14**, e1007807 (2018).
18. T. Lacombe *et al.*, *Theor. Appl. Genet.* **126**, 401–414 (2013).
19. R. Bacilieri *et al.*, *BMC Plant Biol.* **13**, 25–25 (2013).
20. F. Mercati *et al.*, *Front. Plant Sci.* **12**, 692661 (2021).
21. R. Hosfield, J. Cole, *Quat. Sci. Rev.* **190**, 148–160 (2018).
22. A. Timmermann *et al.*, *Nature* **604**, 495–501 (2022).
23. E. C. Corrick *et al.*, *Science* **369**, 963–969 (2020).
24. M. Engel *et al.*, *Quat. Int.* **266**, 131–141 (2012).
25. C. J. Stevens *et al.*, *Holocene* **26**, 1541–1555 (2016).
26. I. Lazaridis *et al.*, *Nature* **536**, 419–424 (2016).
27. C.-C. Wang *et al.*, *Nat. Commun.* **10**, 590 (2019).
28. R. Pinhasi, J. Fort, A. J. Ammerman, *PLOS Biol.* **3**, e410 (2005).
29. I. Mathieson *et al.*, *Nature* **555**, 197–203 (2018).
30. R. Fregel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6774–6779 (2018).
31. I. Olalde *et al.*, *Science* **363**, 1230–1234 (2019).
32. S. Brunel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 12791–12798 (2020).
33. C. Zou *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2023548118 (2021).
34. F. Emanuelli *et al.*, *BMC Plant Biol.* **10**, 241–241 (2010).
35. S. Kobayashi, N. Goto-Yamamoto, H. Hirochika, *Science* **304**, 982–982 (2004).
36. A. R. Walker *et al.*, *Plant J.* **49**, 772–785 (2007).
37. A. R. Walker, E. Lee, S. P. Robinson, *Plant Mol. Biol.* **62**, 623–635 (2006).
38. P. J. Richerson, R. Boyd, R. L. Bettinger, *Am. Antiq.* **66**, 387–411 (2001).
39. R. G. Allaby, C. J. Stevens, L. Kistler, D. Q. Fuller, *Trends Ecol. Evol.* **37**, 268–279 (2022).
40. R. S. Meyer, M. D. Purugganan, *Nat. Rev. Genet.* **14**, 840–852 (2013).
41. Code for: Y. Dong *et al.*, Dual domestications and origin of traits in grapevine evolution, Zenodo (2023); https://doi.org/10.5281/zenodo.7523647.

ACKNOWLEDGMENTS

We thank F. Pelsy, L. Garmendia Auckenthaler, A.-F. Adam-Blondon, C. Cornier, P. Kozma, O. Bachmann, F. Gillet, J.-M. Gobat, S. Dedet, J. Daumann, K. Huber, V. Risovannaya, A. Polulyah, B. Louis, M. Lafargue, G. Jean-Pascal, G. Melyan, D. I. Sumedrea, A. Naqinezhad, M. Filipova, technical staff from EGFV and UEVB, and the Danube-Auen National Park for assistance in the sample collection and laboratory work and P. Kupfer, E. D. O. Roberson, and D. Petkova for comments on the manuscript. **Funding:** This work was supported by the Natural Science Foundation of China (grant 32070599 to W.C.); Yunnan Agricultural University (Research Fund A2032002519 to W.C.); China Agriculture Research System of MOF and MARA CARS-29 (S.W.); the Science Committee at the Ministry of SCS (RA 20APP-4E007 to K.M.); Alliance of International Science Organization (ANSO-CR-PP-2020-04-A to K.M.); Ministerio de Ciencia, Innovación y Universidades and Agencia Estatal de Investigación of Spain (RTI2018-094470-R-C21 to R.A.G.); Predoctoral Fellowship PRE2019-088446 (A.R.I.); Israel Ministry of Science and Technology (90-23-020-12 to E.D.); Fondation Giacomi and Swiss National Science Foundation (SNSF 43307 to C.A.); European Regional Fund (KK.05.1.1.02.0010 to G.Z.); Georgian state budget (L.U., K.B., and T.Z.); TUBITAK and Ministry of Agriculture and Forestry of Republic of Türkiye (grant 105G078 to A.E.); and the Israel Science Foundation (551/18 to E.W.). **Author contributions:** Conceptualization:

Y.D., Z.L., S.W., J.S., W.C.; Formal analysis: S.D., Q.X., X.D.; Funding acquisition: W.C.; Investigation: Y.Z., C.M., S.W., S.L., L.T., C.W., D.L., Y.P., J.L., L.Y., X.L., G.X., Z.Y., B.C., Y.W., P.G., M.R., O.R., A.R.I., Y.W., S.Z.; Resources: Z.L., K.M., M.M., S.G., G.Z., P.F.B., T.L., F.R., P.N., K.B., G.D.B., E.D., G.D.L., J.C., C.F.P., R.A.G., C.A., A.E., Z.D., V.K., G.S., N.G., S.D., N.O., P.T., C.M., V.L., A.J., L.U., T.Z., D.M., M.H., G.J., E.K., T.D., F.G., F.M., F.S., J.E.D., A.M.D., D.C., G.M., T.U., C.Ö., K.K., M.X., J.L., M.Z., L.W., S.J., Y.Z., L.S., S.L.; Supervision: Y.D., H.Y., Y.Z., S.W., J.S., W.C.; Validation: All authors participated in the interpretation of the data; Visualization: S.D., Q.X., X.D.; Writing - original draft: Y.D., S.D., Q.X., X.D., W.C.; Writing - review & editing: W.C. with input from all coauthors. **Competing interests:** A.J. is the founder and owner of Historische Rebsorten vineyard. The remaining authors declare no competing interests. **Data and materials availability:** The VS-1 genome assembly is available at the Genome Warehouse in the National Genomics Data Center, China National Center for Bioinformation, under accession numbers CRA006898 and GWHBQCW00000000. The raw resequencing data are available at the Genome Warehouse in the National Genomics Data Center, China National Center for Bioinformation, under accession number CRA006917. The code for the work can be accessed at Zenodo (*41*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add8655
Materials and Methods
Supplementary Text
Figs. S1 to S27
Tables S1 to S35
References (*42–153*)
MDAR Reproducibility Checklist

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

Submitted 11 July 2022; accepted 23 January 2023
10.1126/science.add8655

HEMATOPOIESIS

USB1 is a miRNA deadenylase that regulates hematopoietic development

Ho-Chang Jeong[1,2]†, Siddharth Shukla[3,4]†‡, Wilson Chun Fok[1,2], Thao Ngoc Huynh[3,4], Luis Francisco Zirnberger Batista[1,2]*, Roy Parker[3,4]*

Mutations in the 3′ to 5′ RNA exonuclease USB1 cause hematopoietic failure in poikiloderma with neutropenia (PN). Although USB1 is known to regulate U6 small nuclear RNA maturation, the molecular mechanism underlying PN remains undetermined, as pre-mRNA splicing is unaffected in patients. We generated human embryonic stem cells harboring the PN-associated mutation c.531_delA in USB1 and show that this mutation impairs human hematopoiesis. Dysregulated microRNA (miRNA) levels in USB1 mutants during blood development contribute to hematopoietic failure, because of a failure to remove 3′-end adenylated tails added by PAPD5/7. Modulation of miRNA 3′-end adenylation through genetic or chemical inhibition of PAPD5/7 rescues hematopoiesis in USB1 mutants. This work shows that USB1 acts as a miRNA deadenylase and suggests PAPD5/7 inhibition as a potential therapy for PN.

Poikiloderma with neutropenia (PN) is an autosomal-recessive bone marrow failure (BMF) syndrome with marked clinical overlap with dyskeratosis congenita (DC) (*1*). However, unlike patients with DC, telomeres are not shortened in patients suffering from PN, providing a distinguishable feature for the correct diagnosis of PN (*2*). PN patients harbor homozygous or compound heterozygous mutations in the human gene *C16orf57*, which encodes the conserved 3′ to 5′ RNA exonuclease U6 biogenesis 1 (USB1) (*2*–*5*). USB1 is required for the processing of U6 and U6atac small nuclear RNAs (snRNAs), and some splicing defects are observed when using yeast and zebrafish models of USB1 deficiency (*6*–*10*). However, lymphoblastoid cells from PN patients do not exhibit reduced U6 snRNA levels and have normal pre-mRNA splicing (*8*). These results establish USB1-mediated PN as a singular BMF syndrome, in which the underlying genetic cause has been identified but the molecular mechanisms leading to tissue failure are unknown.

USB1 mutant hESCs have impaired hematopoietic development

To investigate the role of USB1 in a physiological context, we utilized CRISPR-Cas9 to create human embryonic stem cells (hESCs) containing a frequently occurring c.531_del_A loss-of-function mutation in the USB1 gene (hereafter referred to as USB1 mutant) (fig. S1, A and B). These USB1 mutant hESCs have normal karyotype (fig. S1C), have normal growth rate (fig. S1D), are pluripotent (fig. S1E), and display normal telomere length (fig. S1F), indicating that a clinically relevant USB1 mutation is not deleterious in undifferentiated hESCs.

To elucidate the role of USB1 during hematopoiesis, we performed serum-free hematopoietic differentiations (*11*–*15*) to derive hematopoietic progenitor cells from hESCs (Fig. 1A). Gene expression analysis confirmed the efficiency of this protocol, with silencing of pluripotency markers and efficient formation of hematopoietic lineages at the end (day 30) of differentiation (fig. S2A). USB1 mutant cells did not show any impairment during early stages of hematopoietic differentiation, including the formation of mesoderm (day 3; fig. S2B), and $CD34^+/CD43^-$ hemogenic endothelium (HE) populations (day 8; fig. S2C). However, the formation of $CD45^+$ hematopoietic progenitors (day 16) was decreased in USB1 mutant cells compared to wild-type (WT) cells (fig. S2D), and hematopoietic colony potential analysis showed compromised colony formation in USB1 mutant cells (Fig. 1B). Consistent with a role of USB1 in regulating hematopoiesis, USB1 mRNA levels increased about threefold in mature blood cells compared to undifferentiated hESCs (Fig. 1C). These observations indicate that loss-of-function mutations in USB1 negatively influence hematopoiesis.

As PN is usually associated with severe noncyclic neutropenia (*1*), we specifically analyzed the potential of neutrophil formation in WT and USB1 mutant cells. USB1 mutants had reduced formation of $CD15^+/CD66b^+$ lineages, indicating abnormal neutrophil development (Fig. 1D and fig. S2E). The conditional expression of the WT USB1 protein in USB1 mutants with the use of a Dox-inducible system (fig. S2F) rescued the hematopoietic potential of these cells (Fig. 1E). These results recapitulate major clinical manifestations of USB1 deficiency and establish USB1 as an important regulator of hematopoiesis.

Mutations in USB1 lead to increased 3′ adenylation but do not affect levels of the U6 snRNA in human stem cells and hematopoietic progenitors

To determine the mechanism by which USB1 regulates hematopoiesis, we initially examined if the USB1 mutation affected U6 snRNA. Northern blot analysis of WT and USB1 mutant cells at undifferentiated (D0) and hematopoietic progenitor (D16) stages showed no reduction in the levels of U6 and U6atac snRNAs in USB1 mutants (Fig. 1F and fig. S2, G and H). However, we observed that U6 and U6atac snRNA from USB1 mutant cells were slightly longer compared to WT cells (Fig. 1F), indicating aberrant posttranscriptional processing of these snRNAs similar to what is observed in patient-derived cells (*8*).

Sequencing the 3′ end of U6 snRNA from WT and USB1 mutant cells revealed two changes. First, whereas ~40% of all U6 reads terminated at the +1U adjacent to the Lsm 2-8 boundary site in WT cells, in USB1 mutant cells, additional Us were present, with the majority of ends terminating at the +3U and +4U (day 0: Fig. 1G; day 16: fig. S3A). Similar results were observed for the U6atac snRNA (fig. S3B). Second, we observed that ~25% of extended U6 and U6atac 3′ ends were oligoadenylated in USB1 mutant cells (Fig. 1H and fig. S3, C and D), in contrast to ~5% of 3′ ends of U6 being oligoadenylated in WT cells (Fig. 1G, inset, and fig. S3, E and F). This suggests that the U6 3′ ends are adenylated during U6 snRNA maturation and require USB1 for deadenylation and trimming of the uridylated tail. However, USB1 mutant cells have similar levels of these snRNAs compared to WT cells (Fig. 1F and fig. S2, G and H) and do not exhibit global pre-mRNA splicing changes (see next section), suggesting that USB1 alters other RNAs to affect hematopoiesis.

USB1 mutations affect mRNA levels during blood development, without impairing splicing

To identify other RNAs affected by USB1, we sequenced the transcriptome and miRNome of WT and USB1 mutant cells in the undifferentiated, hematopoietic progenitor ($CD34^+/CD45^+$) and mature blood cell populations (Fig. 1A). We observed few changes in gene expression in USB1 mutants in undifferentiated hESCs (469 out of 15,913 genes were affected with a false discovery rate (FDR) <0.1; 164 genes downregulated more than twofold and 61 genes upregulated more than twofold) (Fig. 2A). As differentiation progressed, we observed more changes in $CD34^+/CD45^+$ cells, suggesting that the defect in hematopoiesis occurs during specific stages of differentiation (3310 genes out of 15,559 genes were affected with an FDR <0.1;

[1]Division of Hematology, Department of Medicine, Washington University in St. Louis, St. Louis, MO 63110, USA. [2]Center for Genome Integrity, Siteman Cancer Center, Washington University in St. Louis, St. Louis, MO 63110, USA. [3]Department of Biochemistry, University of Colorado, Boulder, CO 80303, USA. [4]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA.
*Corresponding author. Email: lbatista@wustl.edu (L.F.Z.B.); roy.parker@colorado.edu (R.P.)
†These authors contributed equally to this work.
‡Present address: Alltrna Inc., Cambridge, MA 02139, USA.

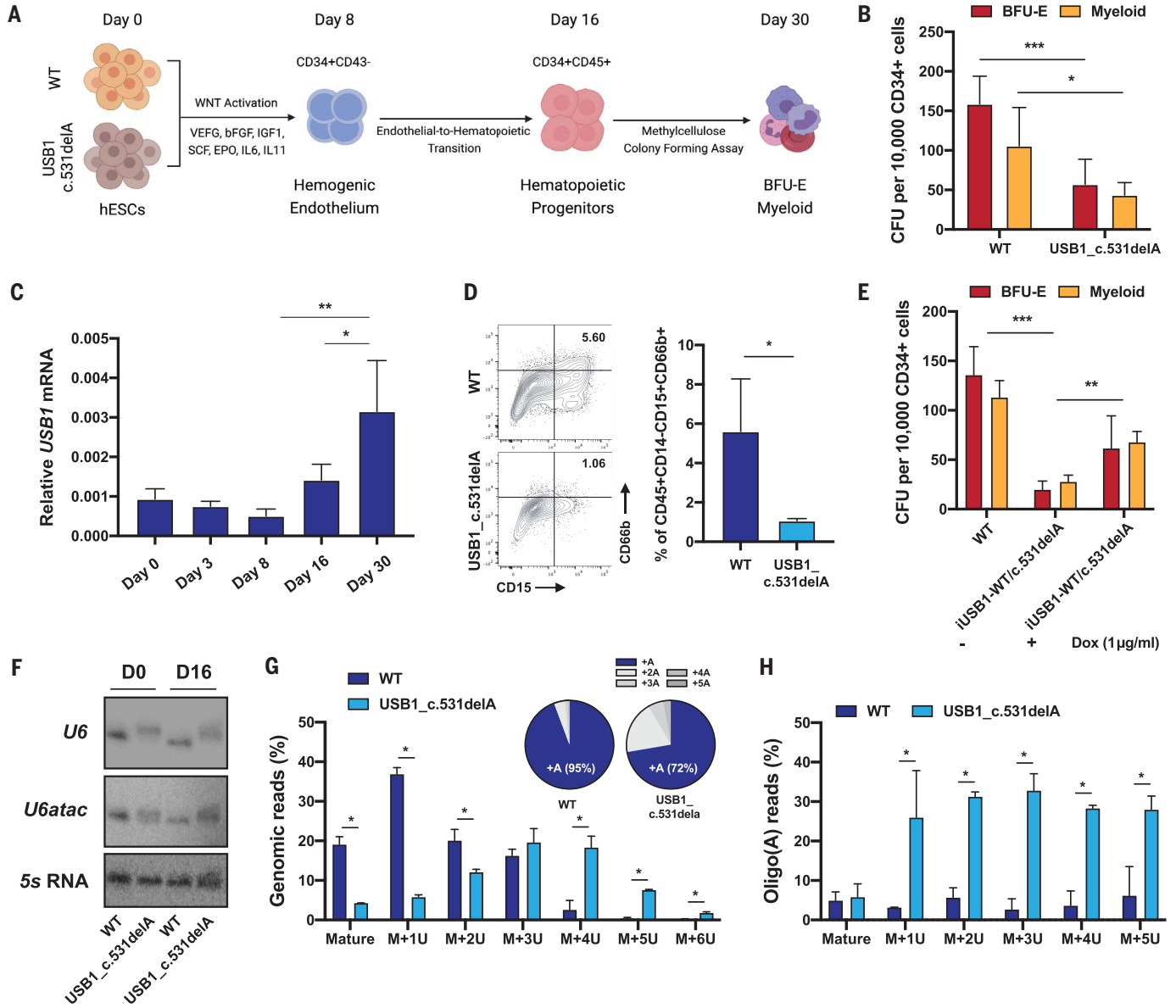


Fig. 1. Loss-of-function mutation in USB1 causes hematopoietic impairment. (**A**) Model depicting the workflow of hematopoietic differentiation from WT or USB1 c.531delA hESCs. Cellular identity is confirmed by expression of the correct differentiation markers at the different steps depicted in the model. (**B**) Colony-forming-cell (CFC) potential of definitive hematopoietic progenitors in WT and USB1 c.531delA cells (mean ± SD, *n* = 6 biological replicates). CFU, colony-forming units. (**C**) USB1 expression levels analyzed at different stages of hematopoietic specification (mean ± SD, *n* = 3 biological replicates). (**D**) Representative flow cytometry analysis for CD15 and CD66b within $CD45^{+}CD14^{-}$ population on day 30 of differentiation (left). The populations were quantified and graphically presented (right, mean ± SD, *n* = 3 biological replicates). (**E**) CFC potential of hematopoietic progenitors in WT and iUSB1-WT/c.531delA cells, treated or not with doxycycline (Dox). Dox (1 µg/ml) was added from day 8 to day 16 of differentiation (mean ± SD, *n* = 3 biological replicates). (**F**) Representative Northern blot for U6 or U6atac snRNA in either undifferentiated hESCs (D0) or $CD34^{+}CD45^{+}$ cells (D16). (**G** and **H**) Bar plots depicting fractions of (G) genomic and (H) posttranscriptionally adenylated U6 snRNA 3′ ends in WT and USB1 c.531delA cells (mean ± SD, *n* = 2 biological replicates). (G) Inset: Length distribution of (A) tails at U6 3′ ends in WT and USB1 c.531delA cells. **P* < 0.05, ***P* < 0.01, and ****P* < 0.001.

704 genes down-regulated more than twofold and 577 genes up-regulated more than twofold) (Fig. 2B). Differentially expressed genes in USB1 mutant cells were enriched for Gene Ontology (GO) pathways involved in regulating cell death and neutrophil differentiation (fig. S4A). Consistent with an effect of USB1 deficiency on neutrophil differentiation (Fig. 1D), we observed a decrease in populations with high granular side scatter in USB1 mutant cells upon direct neutrophil development, when compared to WT cells (fig. S4B). Finally, we also observed more gene expression changes in the USB1 mutant cells when compared to WT cells in their mature blood population (2666 out of 16,310 genes were affected with an FDR <0.1; 510 genes down-regulated more than twofold and 883 genes up-regulated more than twofold) (fig. S4C). Taken together, this suggests that mutations in USB1 have a greater effect on gene expression in differentiating hematopoietic progenitors and mature blood cells, which correlates with an increase in USB1 levels at these specific stages of differentiation (Fig. 1C).

We did not detect global splicing changes in the transcriptome of USB1 mutants at any

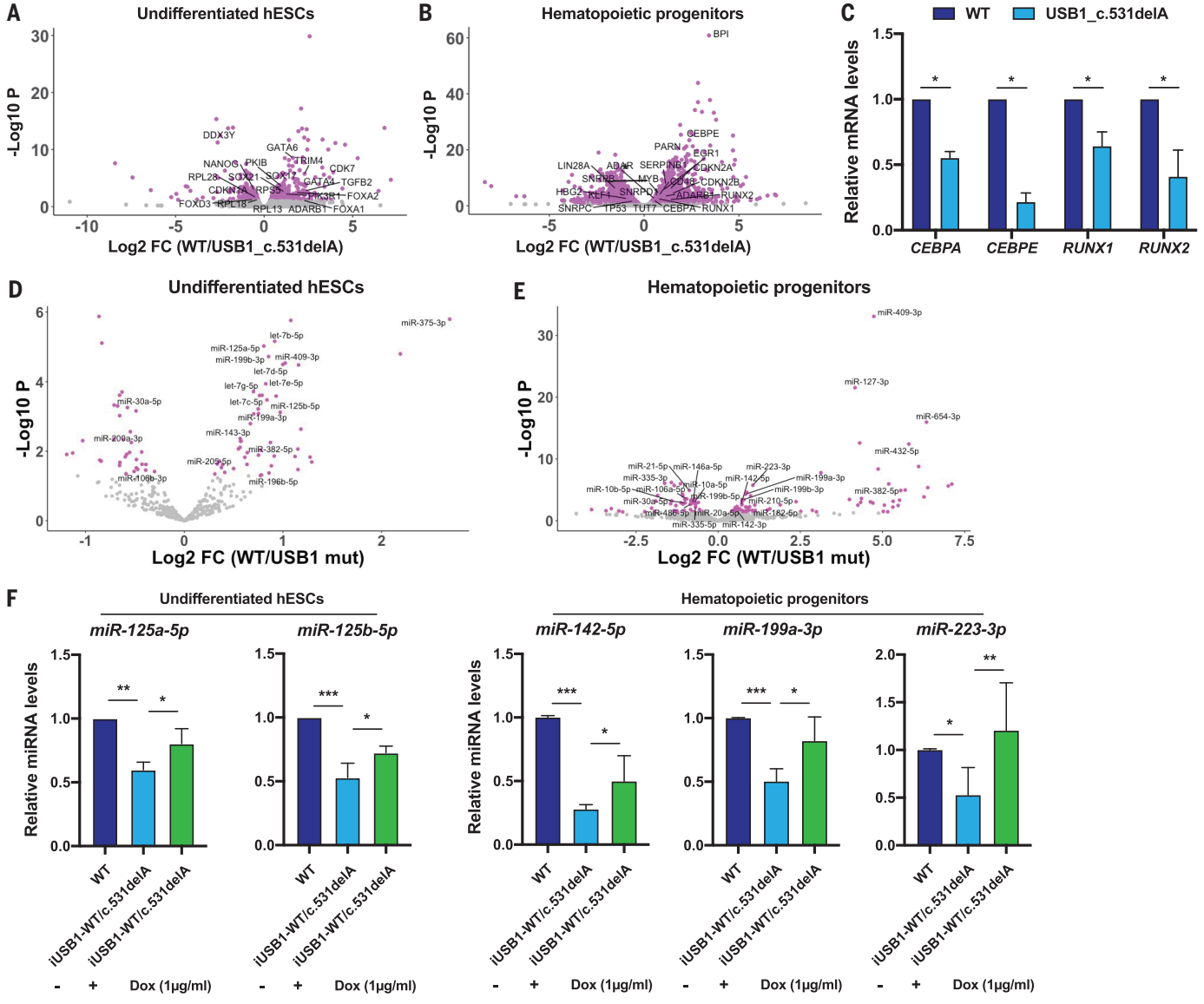


Fig. 2. USB1 mutation affects both mRNA and miRNA levels during hematopoietic differentiation. (**A** and **B**) Volcano plots depicting transcriptome changes in (A) undifferentiated hESCs and (B) $CD34^+CD45^+$ hematopoietic progenitors in WT and USB1 c.531delA cells (gray: unchanged; magenta: differentially expressed). (**C**) Bar plot depicting levels of indicated mRNAs normalized to 5S ribosomal RNA (rRNA) in $CD34^+CD45^+$ hematopoietic progenitors (mean ± SD, n = 3 biological replicates). (**D** and **E**) Volcano plots depicting miRNA changes in (D) undifferentiated hESCs and (E) $CD34^+CD45^+$ hematopoietic progenitors in WT and USB1 c.531delA cells (gray: unchanged; magenta: differentially expressed). (**F**) Bar plot depicting levels of indicated miRNAs as quantified by Mir-X quantitative reverse transcription–polymerase chain reaction (qRT-PCR) normalized to 5S rRNA in both undifferentiated hESCs and hematopoietic progenitors of iUSB1-WT/c.531delA (mean ± SD, n = 3 biological replicates). *P < 0.05, **P < 0.01, and ***P < 0.001.

stage when compared to WT cells, and very few differentially expressed genes were mis-spliced in USB1 mutant cells (fig. S4, D and E). Previous analysis of PN patient lymphoblasts also did not reveal global splicing changes (*8*). This is consistent with normal levels of U6 snRNA and U6atac snRNA in USB1 mutant cells and suggests that USB1 deficiency affects gene expression through a mechanism distinct from pre-mRNA splicing. Next, we investigated whether specific mRNA changes in USB1 mutant cells could explain the hematopoietic failure observed and identified that transcription factors associated with efficient neutrophil formation (i.e., *CEBPA* and *CEBPE*) and overall hematopoiesis (i.e., *RUNX1* and *RUNX2*) are down-regulated in USB1 mutants (Fig. 2C) (*16*–*19*). Combined, these results indicate that despite not affecting U6 levels and pre-mRNA splicing, mutations in USB1 impair correct activation of key hematopoietic and neutrophil development pathways, which is consistent with the reduced output of these populations during development (Fig. 1, D and E).

miRNA levels are progressively altered during targeted hematopoietic development of USB1 mutants

Given that miRNAs can regulate hematopoiesis (*20*), we also investigated whether mutations in USB1 affected miRNA levels in hESCs and their hematopoietic progeny. We observed that the USB1 mutation affected the levels of 82 out of 374 miRNAs in undifferentiated hESCs (P < 0.05; 49 miRNAs were down-regulated and 33 miRNAs were up-regulated) (Fig. 2D). We observed increased changes in miRNA levels

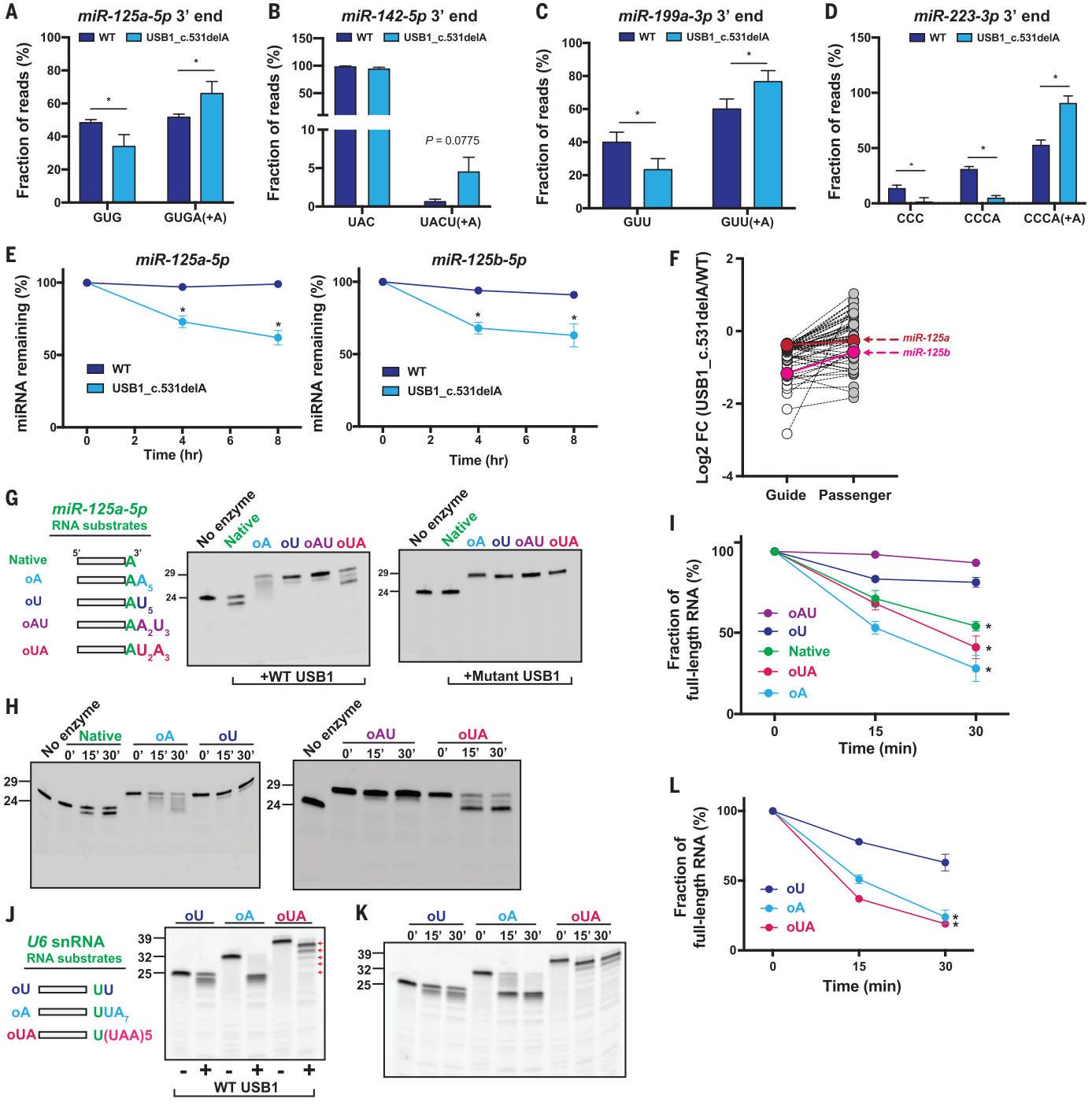


Fig. 3. USB1 deadenylates RNA substrates in vivo and in vitro. (**A**) Bar plot depicting fraction of genomic and posttranscriptionally adenylated miR-125a-5p 3′ ends in undifferentiated WT and USB1 c.531delA cells (mean ± SD, n = 2 biological replicates). (**B** to **D**) Bar plots depicting fraction of genomic and posttranscriptionally adenylated (B) miR-142-5p, (C) miR-199a-3p, and (D) miR-223-3p 3′ ends in CD34^{+}CD45^{+} hematopoietic progenitors in WT and USB1 c.531delA cells (mean ± SD, n = 3 biological replicates). (**E**) Quantification of miR-125a-5p or miR-125b-5p decay rates in WT and USB1 c.531delA at 0, 4, and 8 hours after transcription shutoff (mean ± SD, n = 3 biological replicates). (**F**) Influence of USB1 on miRNA guide and passenger strands. Shown are the expression levels of guide and passenger strands for miRNAs with the highest differential expression between USB1 c.531delA mutants and WT hESCs (mean, n = 3 biological replicates). (**G**) Representative gel images showing processing of miR-125a-5p substrates by WT USB1 or H208Q catalytic mutant. miR-125a-5p sequence includes a genomically encoded A at the 3′ end (shown in green). Marker for 24-nucleotide indicates a native form of miR-125a-5p. (**H**) Representative gel images depicting time-course measurement of USB1's activity on indicated miR-125a-5p RNA substrates. (**I**) Line plot depicting degradation of miR-125a-5p RNA substrates with indicated 3′-end modifications (mean ± SD, n = 3 technical replicates). (**J**) Representative gel image showing processing of U6 snRNA substrates by WT USB1. Genomically encoded bases are shown. Red arrows show position of trimmed UAA intermediates for the oUA substrate incubated with WT USB1. (**K**) Representative gel image depicting time-course measurement of USB1's activity on indicated U6 snRNA substrates. (**L**) Line plot depicting degradation of U6 snRNA substrates with indicated 3′-end modifications (mean ± SD, n = 3 technical replicates). *P < 0.05.

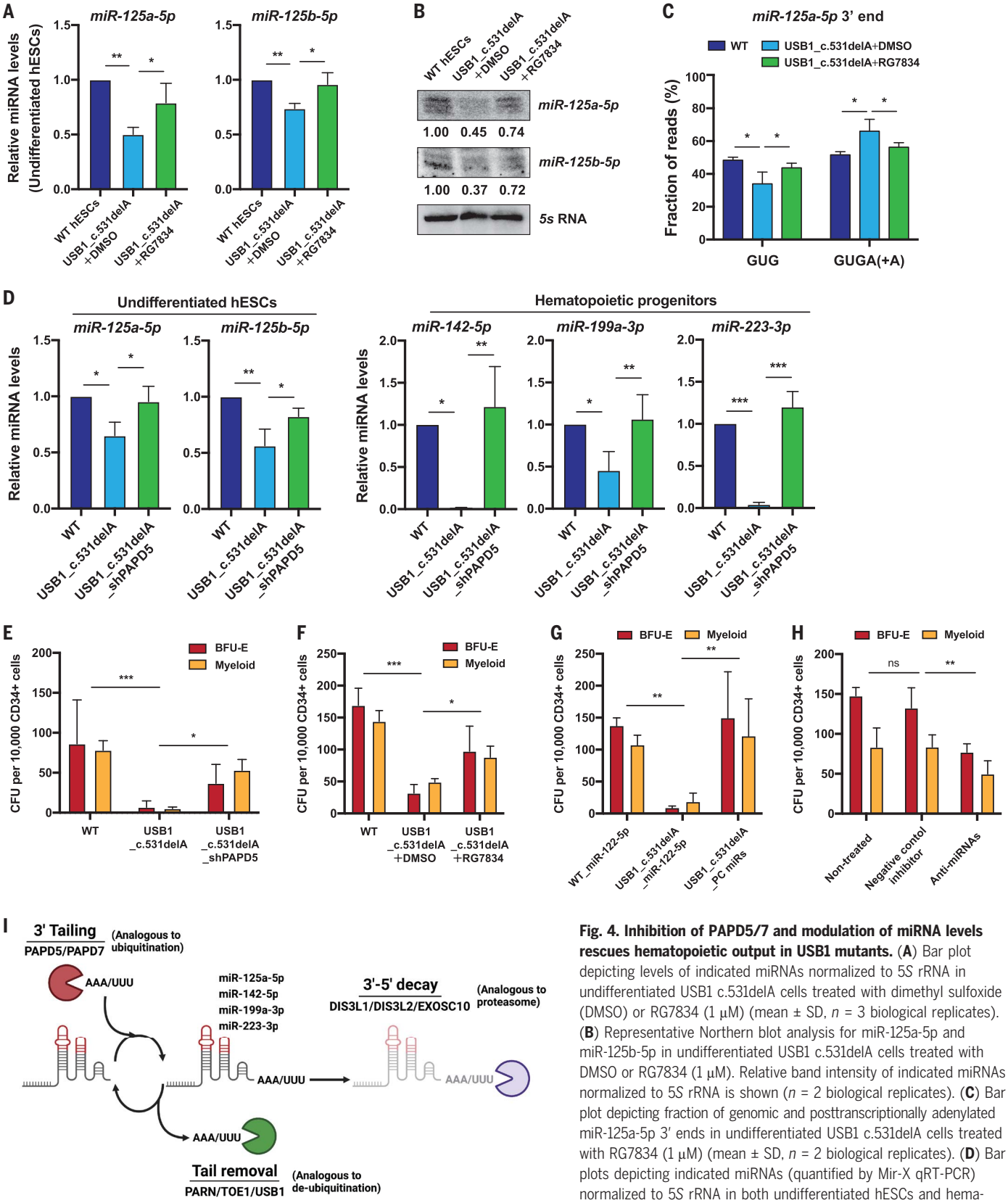


Fig. 4. Inhibition of PAPD5/7 and modulation of miRNA levels rescues hematopoietic output in USB1 mutants. (**A**) Bar plot depicting levels of indicated miRNAs normalized to 5*S* rRNA in undifferentiated USB1 c.531delA cells treated with dimethyl sulfoxide (DMSO) or RG7834 (1 μM) (mean ± SD, *n* = 3 biological replicates). (**B**) Representative Northern blot analysis for miR-125a-5p and miR-125b-5p in undifferentiated USB1 c.531delA cells treated with DMSO or RG7834 (1 μM). Relative band intensity of indicated miRNAs normalized to 5*S* rRNA is shown (*n* = 2 biological replicates). (**C**) Bar plot depicting fraction of genomic and posttranscriptionally adenylated miR-125a-5p 3′ ends in undifferentiated USB1 c.531delA cells treated with RG7834 (1 μM) (mean ± SD, *n* = 2 biological replicates). (**D**) Bar plots depicting indicated miRNAs (quantified by Mir-X qRT-PCR) normalized to 5*S* rRNA in both undifferentiated hESCs and hematopoietic progenitors in USB1 c.531delA_shPAPD5 cells (mean ± SD, *n* = 3 biological replicates). (**E**) CFC potential of definitive hematopoietic progenitors in WT, USB1 c.531delA, and USB1 c.531delA_shPAPD5 cells (mean ± SD, *n* = 3 biological replicates). (**F**) CFC potential of

definitive hematopoietic progenitors in WT and USB1 c.531delA treated with DMSO or RG7834 (mean ± SD, *n* = 3 biological replicates). RG7834 (0.5 μM) was added from day 8 to day 16 of differentiation every other day. (**G**) CFC potential of definitive hematopoietic progenitors in WT_miR-122-5p, USB1_miR-122-5p, and USB1 c.531delA_PC miRs. PC miRs indicate the polycistronic expression of miR-125a-5p, 142-5p, 199a-3p, and 223-3p. The expression of miR-122-5p in WT and USB1 mutant cells was performed as a control (mean ± SD, *n* = 3 biological replicates). (**H**) CFC potential of definitive hematopoietic progenitors in WT cells treated with miRNA inhibitors (Anti-miRs) specifically targeting miR-125a-5p, 142-5p, 199a-3p, and 223-3p. miRNA inhibitors (20 nM) were added from day 8 to day 16 of differentiation (mean ± SD, *n* = 4 biological replicates). (**I**) Model depicting the regulation of noncoding RNA stability through competition between "Tailing" and "Tail removal" enzymes. This system is analogous to the ubiquitin-mediated proteasome degradation system, in which RNAs are tagged for degradation by 3′-end modification by enzymes such as PAPD5 and PAPD7, and protector exonucleases such as PARN and USB1 remove the posttranscriptional modifications to stabilize the RNA. In the absence of tail removal, the 3′-end–modified RNA would be degraded by 3′ to 5′ exonucleases such as EXOSC10. **P* < 0.05, ***P* < 0.01, and ****P* < 0.001. ns, not significant.

in $CD34^+/CD45^+$ hematopoietic cells in USB1 mutants (131 out of 771 miRNAs were affected with $P < 0.05$; 82 miRNAs were down-regulated and 49 miRNAs were up-regulated) (Fig. 2E). KEGG analysis showed that differentially expressed miRNAs in USB1 mutant cells predominantly affect pathways involved in cancer progression, including acute myeloid leukemia, which is frequently associated with PN (fig. S4F) (*21*). We verified the down-regulation of specific miRNAs in USB1 mutant cells and found that miR-125a-5p, miR-125b-5p, miR-142-5p, miR-199a-3p, and miR-223-3p [which are predominately involved in erythroid, myeloid, and granulocytic differentiation (*20*, *22*)] were decreased in USB1 mutant cells compared to WT cells at different stages of hematopoietic development (Fig. 2F). Expression of the WT USB1 protein with a Dox-inducible system rescued the levels of these different miRNAs during different stages of hematopoietic development (Fig. 2F). In contrast to changes in miRNA and mRNA levels, we observed few significant changes in other noncoding RNAs (ncRNAs) in undifferentiated (fig. S4, G and H) or $CD34^+/CD45^+$ (fig. S4I) USB1 mutant cells. These data suggest that USB1 regulates hematopoietic development via miRNA homeostasis, although we cannot rule out the formal hypothesis that other types of ncRNAs might also be involved.

USB1 deadenylates miRNAs and regulates their stability

We hypothesized that USB1 might regulate miRNAs by removing 3′-end adenylated tails that would otherwise trigger miRNA degradation for three reasons. First, 3′ oligoadenylation of miRNAs by the noncanonical polyadenylate [poly(A)] polymerases PAPD5 and PAPD7 can promote their degradation by the cytoplasmic 3′ to 5′ exonucleases DIS3L and/or DIS3L2 (*23*). Second, removal of oligo(A) tails by the poly(A)-specific nuclease PARN stabilizes miRNAs (*23*). Finally, USB1, although generally thought to act on U tails, can remove poly(A) tails in vitro (*24*). This hypothesis predicts that miRNAs regulated by USB1 would show increased levels of 3′ adenylation in USB1 mutant cells and that recombinant USB1 protein would deadenylate 3′-adenylated miRNAs.

To test this hypothesis, we initially sequenced the 3′ ends of four miRNAs that were reduced in USB1 mutant cells and known to regulate hematopoiesis (*20*, *22*). The USB1 mutation led to an increase in the levels of adenylated reads at 3′ ends of miR-125a-5p in undifferentiated hESCs (Fig. 3A). In WT cells, miR-125a-5p terminates at either the genomically encoded 5′-GUG-3′ or at 5′-GUGA-3′ in a 1:1 ratio (Fig. 3A). However, in USB1 mutant cells, we observed a decrease in the 5′-GUG-3′ fraction and a proportional increase in the 5′-GUGA-3′ fraction (Fig. 3A). Similar results were observed for miR-142-5p, miR-199a-3p, and miR-223-3p in $CD34^+/CD45^+$ hematopoietic progenitors, with USB1 mutants containing more adenylated ends when compared to WT (Fig. 3, B to D). Broader analysis of miRNA sequencing in WT and USB1 mutant hESCs (fig. S5A), or in WT and K562 hematopoietic cells in which we ablated USB1 (USB1-KO; fig. S5B), demonstrated that many miRNAs with reduced expression levels show increased 3′ adenylation in USB1 mutants, either in hESCs or K562s. A possibility that at this point cannot be ruled out is that USB1 might also increase levels of some miRNAs by steric inhibition of nucleases, and/or by creating a 2′-3′ cyclic phosphate at the 3′ end of miRNAs, a USB1-dependent modification shown in the U6 snRNA (*8*, *25*).

Two observations provide evidence that USB1 affects the decay rate of some miRNAs. First, after transcriptional inhibition with actinomycin D, USB1 mutant hESCs show increased rates of decay of miR-125a-5p and miR-125b-5p (Fig. 3E) and miR-142-5p, miR-199a-3p, and miR-223-3p (fig. S5C). These results were further confirmed when we compared miRNA decay rates in WT and USB1-KO K562 hematopoietic cells (fig. S5D). Second, although guide miRNA strands decreased, we observed that levels of passenger miRNA strands remained similar in WT and USB1 mutants (fig. S5C). These results were also observed in K562 cells (fig. S5E) and further confirmed by miRNA sequencing analysis (fig. S6A), which shows that at a global level, USB1 impairment mostly affects guide, and not passenger miRNA strands (Fig. 3F: hESCs; fig. S6B: K562). In addition, expression of WT USB1 did not affect levels of passenger miRNA strands in USB1 mutant hESCs (fig. S6C).

To assess whether USB1 could directly deadenylate an adenylated miRNA, we purified recombinant human USB1 (WT and a catalytically inactive mutant H208Q) and tested its activity on 5′-5(6)-carboxyfluorescein–labeled miR-125a-5p substrates with different 3′-end additions (Fig. 3G). We observed that WT, but not catalytically inactive USB1, efficiently removed the adenosine(s) from the 3′ end of miR-125a-5p in a time-dependent manner, as seen by the shortened products observed for native miR-125a-5p (which has a single A at its 3′ end), oA and oUA substrates (Fig. 3, G to I). USB1 was also able to rapidly remove adenosines from the 3′ end of a U6 snRNA oligonucleotide (Fig. 3, J to L). This demonstrates that USB1 can function as a deadenylase for miRNAs and U6 snRNAs in vitro.

Our data suggest that USB1 can also remove single uridine residues from RNA substrates but cannot efficiently remove poly(U) tails. Specifically, USB1 slowly removed one uridine from the 3′ end of oligouridylated miR-125a-5p substrates, as observed by the appearance of a small amount of product shorter by one nucleotide (Fig. 3, G to I), which we confirmed with shorter RNA substrates (fig. S6, D and E). USB1 also removed a single uridine residue from an oligouridylated 3′ end of the U6 snRNA (Fig. 3, J and K).

Modulation of miRNA 3′ adenylation by PAPD5 rescues miRNA levels and restores hematopoietic output in USB1 mutants

Our data indicate that a loss of expression of different miRNAs required for blood development contributes to the hematopoietic deficit observed in PN patients owing to the failure of USB1 to remove 3′-end adenylated tails. Accordingly, we observed that inhibiting PAPD5/7 noncanonical poly(A) polymerases with RG7834 (*26*) rescued levels of miRNAs that were reduced in USB1 mutant cells (Fig. 4, A and B). Moreover, treatment of USB1 mutants with RG7834 led to a decrease in adenylated miR-125a-5p, which was compensated by an increase in the nonadenylated form, demonstrating that these enzymes are responsible for adding adenylated tails to the 3′ end of miRNAs that are increased in USB1 mutants (Fig. 4C). Inhibition of PAPD5 by constitutive silencing using short hairpin RNAs (fig. S7A) also rescued the

levels of miRNAs affected by USB1 mutation to approximately those of the WT (Fig. 4D). We did not observe any changes in the levels of passenger miRNA strands after chemical inhibition or genetic silencing of PAPD5/7 (fig. S7, B and C).

These results suggested that inhibition of PAPD5/7 activity should rescue the hematopoietic development defects seen in USB1 mutant cells by preventing the 3′-end adenylation of miRNAs. To test this hypothesis, we investigated whether the genetic silencing of PAPD5 or treatment with RG7834 could rescue hematopoietic differentiation of USB1 mutants. We observed that the compromised colony-formation potential in USB1 mutants was rescued by both the genetic (Fig. 4E) and chemical (Fig. 4F) inhibition of PAPD5. The chemical inhibition of PAPD5/7 by RG7834 treatment in WT cells did not affect miRNA levels (fig. S7D) or colony-forming potential (fig. S7E), consistent with our previous report (*14*). RG7834 treatment also improved neutrophil formation in USB1 mutant cells, a key feature of hematopoietic failure in PN (fig. S7F). These results indicate that inhibition of PAPD5/7 could be a potential therapeutic strategy for treating these patients. Treatment with RG7834 did not rescue the extended U6 3′ end (fig. S7G), further supporting hematopoietic failure in PN as being independent from U6.

Rescue of miRNA levels restores hematopoietic development in USB1 mutants

To directly test whether the miRNAs downregulated in USB1 mutants are responsible for the hematopoietic failure observed, we engineered USB1_c.531delA mutant hESCs with constitutive expression of miR-125a-5p, miR-142-5p, miR-199a-3p, and miR-223-3p (fig. S7H). This led to a rescue of both erythroid and myeloid hematopoietic outputs in USB1 mutants, to levels similar to those of WT cells (Fig. 4G). This demonstrates that miRNA deficiencies are linked to hematopoietic failure in USB1 mutants. Moreover, treating WT $CD34^+$ hematopoietic cells with miRNA inhibitors targeting miR-125a-5p, miR-142-5p, miR-199a-3p, and miR-223-3p (fig. S7I) caused a significant decline in their colony-forming ability (Fig. 4H). These results demonstrate that the hematopoietic failure observed in USB1 mutants is directly linked to impaired levels of these miRNAs. This conclusion is corroborated by the role of Dicer and Ago-2 in blood cell development (*27*–*29*).

Conclusion

In conclusion, we have shown that USB1 functions to deadenylate miRNAs, limiting their degradation and increasing their abundance, and this deadenylating activity of USB1 regulates hematopoiesis. This identifies USB1 as a second de-tailing enzyme, which, similar to PARN (*23*), can remove oligo(A) tails from ncRNAs to enhance their stability (Fig. 4I). Inhibition of the enzymes responsible for miRNA adenylation, PAPD5 and PAPD7, rescued the hematopoietic deficit observed in the USB1 mutant cells. Our results provide a molecular understanding of the pathogenesis observed in patients with USB1 mutations and suggest that PAPD5/7 inhibitors might be a treatment option for PN patients.

REFERENCES AND NOTES

1. A. J. Walne *et al*., *Haematologica* **101**, 1180–1189 (2016).
2. A. J. Walne, T. Vulliamy, R. Beswick, M. Kirwan, I. Dokal, *Hum. Mol. Genet.* **19**, 4453–4461 (2010).
3. D. Concolino *et al*., *Am. J. Med. Genet. A.* **152A**, 2588–2594 (2010).
4. A. Tanaka *et al*., *Am. J. Med. Genet. A.* **152A**, 1347–1348 (2010).
5. L. Volpi *et al*., *Am. J. Hum. Genet.* **86**, 72–76 (2010).
6. S. Mroczek *et al*., *Genes Dev.* **26**, 1911–1925 (2012).
7. V. Shchepachev, H. Wischnewski, E. Missiaglia, C. Soneson, C. M. Azzalin, *Cell Rep.* **2**, 855–865 (2012).
8. C. Hilcenko *et al*., *Blood* **121**, 1028–1038 (2013).
9. E. A. Colombo *et al*., *Sci. Rep.* **5**, 15814 (2015).
10. V. Shchepachev, H. Wischnewski, C. Soneson, A. W. Arnold, C. M. Azzalin, *FEBS Lett.* **589**, 2417–2423 (2015).
11. C. M. Sturgeon, A. Ditadi, G. Awong, M. Kennedy, G. Keller, *Nat. Biotechnol.* **32**, 554–561 (2014).
12. A. Ditadi, C. M. Sturgeon, *Methods* **101**, 65–72 (2016).
13. W. C. Fok *et al*., *Stem Cell Reports* **9**, 409–418 (2017).
14. W. C. Fok *et al*., *Blood* **133**, 1308–1312 (2019).
15. S. Shukla, H. C. Jeong, C. M. Sturgeon, R. Parker, L. F. Z. Batista, *Blood Adv.* **4**, 2717–2722 (2020).
16. R. Morosetti *et al*., *Blood* **90**, 2591–2600 (1997).
17. R. Yamanaka *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13187–13192 (1997).
18. R. Avellino, R. Delwel, *Blood* **129**, 2083–2091 (2017).
19. M. de Bruijn, E. Dzierzak, *Blood* **129**, 2061–2069 (2017).
20. C. Z. Chen, L. Li, H. F. Lodish, D. P. Bartel, *Science* **303**, 83–86 (2004).
21. V. Licursi, F. Conte, G. Fiscon, P. Paci, *BMC Bioinformatics* **20**, 545 (2019).
22. F. Fazi *et al*., *Cell* **123**, 819–831 (2005).
23. S. Shukla, G. A. Bjerke, D. Muhlrad, R. Yi, R. Parker, *Mol. Cell* **73**, 1204–1216.e4 (2019).
24. Y. Nomura, D. Roston, E. J. Montemayor, Q. Cui, S. E. Butcher, *Nucleic Acids Res.* **46**, 11488–11501 (2018).
25. A. L. Didychuk *et al*., *Nat. Commun.* **8**, 497 (2017).
26. H. Mueller *et al*., *Hepatology* **69**, 1398–1411 (2019).
27. D. O'Carroll *et al*., *Genes Dev.* **21**, 1999–2004 (2007).
28. S. Guo *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **107**, 14229–14234 (2010).
29. M. F. Alemdehy *et al*., *Blood* **119**, 4723–4730 (2012).

ACKNOWLEDGMENTS

The authors thank members of the Parker and Batista labs for valuable advice and input. S.S. thanks C. J. Lim and A. Gooding (Tom Cech lab) for assistance with the AKTA EPLC and advice regarding protein purification. R.P., L.B., S.S., and H.J. are inventors on a provisional patent (17/775,133) filed by the University of Colorado Boulder that covers Usb1 as a target in leukemia. **Funding:** Howard Hughes Medical Institute (S.S., T.N.H., R.P.); National Institutes of Health NHLBI (1R01HL137793-01) (L.F.Z.B.); Department of Defense (L.F.Z.B., R.P.); American Cancer Society (L.F.Z.B.); Siteman Cancer Center at WUSTL (L.F.Z.B.); Center of Regenerative Medicine at WUSTL (L.F.Z.B.); National Research Foundation of Korea (NRF-2021R1A6A3A03045808) (H.-C.J.). **Author contributions:** Conceptualization: H.-C.J., S.S., R.P., L.F.Z.B. Methodology: H.-C.J., S.S., W.C.F., T.N.H. Investigation: H.-C.J., S.S., W.C.F. Visualization: H.-C.J., S.S., W.C.F. Funding acquisition: H.-C.J., R.P., L.F.Z.B., Writing – revised draft: H.-C.J., S.S., T.N.H., R.P., L.F.Z.B. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Raw sequencing data reported in this article have been deposited in the Gene Expression Omnibus Database (accession numbers GSE224430; GSE224431; GSE224432; GSE224501).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj8379
Materials and Methods
Figs. S1 to S7
Tables S1 to S3
References (*30*–*39*)
MDAR Reproducibility Checklist

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

Submitted 7 June 2021; resubmitted 4 April 2022
Accepted 3 February 2023
10.1126/science.abj8379

CORRELATED ELECTRONS

Observation of a critical charge mode in a strange metal

Hisao Kobayashi[1,2]*, Yui Sakaguchi[1], Hayato Kitagawa[1,2], Momoko Oura[1,2], Shugo Ikeda[1,2], Kentaro Kuga[3], Shintaro Suzuki[3], Satoru Nakatsuji[3,4,5,6]*, Ryo Masuda[2,7,8], Yasuhiro Kobayashi[2,7], Makoto Seto[2,7], Yoshitaka Yoda[9], Kenji Tamasaku[2], Yashar Komijani[10,11], Premala Chandra[11], Piers Coleman[11,12]*

Understanding the strange metallic behavior that develops at the brink of localization in quantum materials requires probing the underlying electronic charge dynamics. Using synchrotron radiation–based Mössbauer spectroscopy, we studied the charge fluctuations of the strange metal phase of β-$YbAlB_4$ as a function of temperature and pressure. We found that the usual single absorption peak in the Fermi-liquid regime splits into two peaks upon entering the critical regime. We interpret this spectrum as a single nuclear transition, modulated by nearby electronic valence fluctuations whose long time scales are further enhanced by the formation of charged polarons. These critical charge fluctuations may prove to be a distinct signature of strange metals.

The strange metal (SM) is a ubiquitous state of matter found to develop in quantum materials with strong correlations, often appearing as a fan-shaped region of the phase diagram centered around an unstable quantum critical (QC) point. The characteristics of SMs include a logarithmic temperature (T) dependence of specific heat $C/T \sim -\log T$, a linear-in-T resistivity $\rho(T) \sim T$ (*1*), and a strong violation of Kohlers law in the magnetotransport (*2*–*4*). These properties and their universality defy the standard concept of quasiparticle excitations, which is central to the Fermi liquid (FL) theory of metals. This enigma has prompted a wide range of proposals for the origin of SM behavior, including Fermi surface instabilities (*1*, *5*–*7*), valence quantum criticality (*8*), charge stripes (*9*), and nematicity (*10*–*12*); it has also motivated approaches such as holographic duality (*13*–*15*) and simulation by use of cold atoms (*16*).

Although the spin dynamics at quantum criticality has been extensively studied, little is known experimentally about the charge dynamics because appropriate laboratory probes are scarce. Conventionally, charge dynamics are studied with optical spectroscopy (*17*), but these methods probe only the low-momenta, divergence-free transverse components of the current density that, by the continuity equation, do not couple to fluctuations in the charge density. Longitudinal current fluctuations can be probed by means of electron energy loss spectroscopy (EELS) but to date are limited to energies above the Debye energy because of difficulties in subtracting the phonon background in the signal (*18*–*20*). A classic method to detect low-frequency longitudinal charge dynamics is Mössbauer spectroscopy, successfully used in the past to detect the slowing of the charge dynamics at charge-ordering transitions of europium- and iron-based compounds (*21*, *22*).

However, the widespread adoption of Mössbauer methods has long been hindered by the lack of suitable radioisotope sources. To overcome these difficulties, a new generation of Mössbauer spectroscopy has recently been developed by using synchrotron radiation (SR) (*23*). SR-based Mössbauer spectroscopy (Fig. 1A) can be used for a wide range of Mössbauer isotopes, providing improved energy resolution for the isotopes with shorter lifetimes; it offers an unprecedented capability to select a particular nuclear transition, taking advantage of the perfectly polarized SR. This approach presents an ideal probe to resolve longitudinal charge dynamics in materials for which conventional Mössbauer techniques are inapplicable.

We report a direct observation of critical charge dynamics in a SM regime by using SR-based ^{174}Yb Mössbauer spectroscopy. The heavy fermion metal β-$YbAlB_4$ provides an ideal platform to study the SM regime at ambient pressure in a stoichiometric crystal (*3*, *24*). In β-$YbAlB_4$, core level x-ray studies have established the presence of an intermediate valence state caused by valence fluctuations between two ionic configurations: $Yb^{2+} \rightleftharpoons Yb^{3+} + e^-$ (*25*). Usually, in heavy fermion compounds, such valence fluctuations are too fast to be observed with Mössbauer spectroscopy (*26*–*29*), but we show that this is not the case in the SM regime.

Using synchrotron radiation–based Mössbauer spectroscopy to study charge fluctuations

Mössbauer spectroscopy measures the shift in a nuclear absorption line caused by changes in the local (q-integrated) charge density. The characteristic time scale of the measurement is the lifetime of the nuclear excited state, $\tau_0 \sim 2.5$ ns in ^{174}Yb. Charge fluctuations that are much shorter in time than τ_0 produce a single motionally narrowed absorption line, whereas charge fluctuations that are much longer in time than τ_0 produce a double peak absorption line, corresponding to the two different valence states of the Yb ion (Fig. 1C). By fitting the Mössbauer absorption line shape, one can detect charge fluctuations with time scales in the range of ~$0.1\tau_0$ to ~$10\tau_0$ (*30*).

β-$YbAlB_4$ exhibits quantum criticality without tuning in an intermediate valence state (*25*), and the application of an infinitesimal magnetic field B tunes the SM into a FL with $k_B T_{FL} \sim \mu_B B$, where k_B, T_{FL}, and μ_B are the Boltzmann constant, FL temperature, and the Bohr magneton, respectively. The slope of the linear-in-T resistivity $\rho(T) \sim T$ over T between 0.5 and 25 K at ambient pressure corresponds to a nearly quantum-saturated scattering rate $\tau_{tr}^{-1} = 0.4 \times k_B T/\hbar$ (*30*), thus establishing β-$YbAlB_4$ as a system with Planckian dissipation (*31*). This anomalous $\rho(T)$ and its extension over a broad pressure (p) range from ambient pressure to $p^* \sim 0.5$ GPa (Fig. 1B) (*3*, *24*, *32*) provides an excellent setting for high-precision measurements of the critical charge fluctuations, likely of relevance to the broader family of SMs.

Measuring charge dynamics in β-$YbAlB_4$

We investigated how the QC behavior in the SM regime affects the charge dynamics, following their evolution as the SM regime at ambient pressure transforms into a FL regime under pressure. At 20 K and ambient pressure (Fig. 2A), the Mössbauer spectra exhibit a single line feature. However, below $T^* \sim 10$ K, as one enters the QC region, this peak broadens into a two-peak structure, with 5σ significance (*30*). This two-peak structure observed for $p < 0.7$ GPa at 2 K coalesces into a single peak at around $p \sim 1.2$ GPa, ultimately sharpening into an almost resolution-limited peak at $p = 2.3$ GPa, which is characteristic of a FL (Fig. 2B) (*30*).

[1]Graduate School of Material Science, University of Hyogo, 3-2-1 Koto, Hyogo 678-1297, Japan. [2]RIKEN SPring-8 Center, Hyogo 679-5148, Japan. [3]Institute for Solid State Physics, University of Tokyo, Kashiwa 277-8581, Japan. [4]Department of Physics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. [5]Trans-scale Quantum Science Institute, University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan. [6]Institute for Quantum Matter and Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA. [7]Institute for Integrated Radiation and Nuclear Science, Kyoto University, Osaka 590-0494, Japan. [8]Graduate School of Science and Technology, Hirosaki University, Aomori 036-8561 Japan. [9]Japan Synchrotron Radiation Research Institute, Hyogo 679-5198, Japan. [10]Department of Physics, University of Cincinnati, Cincinnati, OH 45221-0011, USA. [11]Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854, USA. [12]Hubbard Theory Consortium, Department of Physics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK.
*Corresponding author. Email: kobayash@sci.u-hyogo.ac.jp (H.K.); satoru@phys.s.u-tokyo.ac.jp (S.N.); coleman@physics.rutgers.edu (P.C.)

The local symmetry at the Yb site of β-$YbAlB_4$ with the orthorhombic structure allows us to rule out a nuclear origin of the double-peak structure. For $c \| k_0$ (the propagation vector of the incident x-ray), the symmetry selects two degenerate nuclear transitions $I_g = 0 \rightarrow I_e^z = \pm 1$ from the five $E2$ nuclear transitions (ΔI^z = 0, ±1, and ±2) of the ^{174}Yb Mössbauer resonance (Fig. 1, A and C) (*33*). The absence of magnetic order in β-$YbAlB_4$ (*24*, *32*) also eliminates as explanations magnetic and nonaxially symmetric quadrupolar hyperfine interactions (*30*). This leaves a combination of the electric monopole and axially symmetric quadrupolar interactions—linking the hyperfine energy to the valence state of the rare-earth ion—as the only candidate for the observed splitting. The presence of a Mössbauer line splitting then implies a distribution of Yb valences within the crystal. We argue that these result from slow dynamic charge fluctuations.

All Yb sites are crystallographically equivalent in β-$YbAlB_4$, and SR x-ray diffraction measurements (*34*) show that the lattice structure does not change up to 3.5 GPa at 7 K; furthermore, the absence of any low-temperature phase transitions rules out the possibility of a charge density wave (*30*). Moreover, the residual resistivity ratio (RRR) exceeds 100, indicating the low levels of quenched disorder in this material. Given that disorder broadens the Mössbauer absorption peak, our ability to resolve the double-peak structure is consistent with this conclusion. An attempt to fit the Mössbauer spectrum with two nuclear transitions (a static hyperfine interaction), using a width corresponding to the experimental energy resolution, fails to reconstruct the feature at 2 K and ~0 mm/s (Fig. 2A, blue dashed line). Thus, the two-peak structure and line broadening observed for $T < 5$ K and $p < 0.7$ GPa must derive from a single nuclear transition that is dynamically modulated by fluctuations between two different Yb charge states (a time-dependent hyperfine interaction) (Fig. 1C) (*30*).

We analyzed our Mössbauer spectra at ambient pressure using a stochastic theory (*35*–*37*) with a single nuclear transition modulated by two different charge states (*30*). The predicted spectra (Fig. 2A, red lines) well reproduce the two-peak structure in the spectra at low T and its subsequent collapse into a single line with increasing T.

At ambient pressure, the extracted fluctuation time τ_f between two different Yb charge states is unusually long compared with the electronic time scales, exhibiting a slow power-law growth $T^{-\eta}$ ($\eta \sim 0.2$) on cooling below T^* (Fig. 2C). The energy difference between two selected nuclear transitions is almost independent of T up to 20 K (*30*), so that the development of the two-peak structure in the observed spectra must derive from the marked low-T growth in τ_f. On the other hand the gradual collapse of the two-peak structure in the observed ^{174}Yb Mössbauer spectra at 2K with increasing p indicates that fluctuation time scale τ_f becomes shorter as p increases (Fig. 2B). The spectra at $p < 1.2$ GPa can be analyzed and reconstructed with the same stochastic model used at the ambient pressure, whereas the spectrum observed at 2.3 GPa was simply fit by using the static model. The linewidth of this single absorption component was found to be $\Gamma = 1.11$ mm/s, which is slightly broader than the resolution limit $\Gamma_0 = \hbar/\tau_0 = 1.00$ mm/s (3 mK), for ^{174}Yb Mössbauer spectroscopy ($\tau_0 = 2.58$ ns, and $\hbar$ is Planck's constant h divided by 2π).

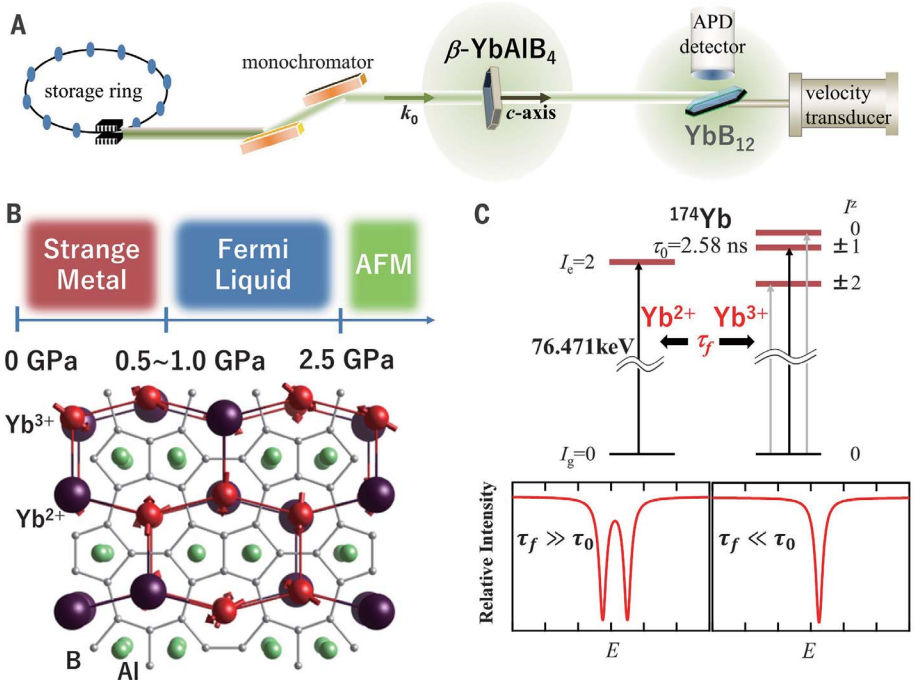


Fig. 1. Experimental setup and concept. (**A**) Schematic of our experimental setup for the synchrotron radiation–based ^{174}Yb Mössbauer spectroscopy (*47*). The ^{174}Yb nuclear resonance (E_γ = 76.471 keV) was obtained from synchrotron radiation by use of a monochromator. The c axis of the single-crystalline β-$YbAlB_4$ samples was aligned along the propagation vector k_0 of the incident x-ray under both ambient and external pressure. The single-crystalline YbB_{12} samples were cooled at 26 K. A Si avalanche photodiode (APD) detector was used to measure delayed incoherent emission from ^{174}Yb nuclei in the YbB_{12}. (**B**) (Top) Schematic phase diagram of β-$YbAlB_4$ as a function of pressure at low temperatures. (Bottom) Illustration of the crystal structure of β-$YbAlB_4$ with a snapshot of the Yb valences, Yb^{2+} (large purple spheres) and Yb^{3+} (small red spheres, with arrows indicating magnetic moment). (**C**) (Top) Energy level diagram of a ^{174}Yb nucleus in Yb^{2+} and Yb^{3+} ions. The lowest excited states of a ^{174}Yb nucleus lie in a I_e = 2 multiplet with a lifetime τ_0 = 2.58 ns. The excited-state energies are perturbed by the electron charge distributions around the nucleus; a spherically symmetric charge distribution (Yb^{2+}) preserves the multiplet degeneracy, whereas one with axial symmetry (Yb^{3+}) splits the multiplet into two doublets and a singlet. The allowed Mössbauer transitions are indicated with arrows, where the black arrows represent the two selected transitions for photons travelling along the c axis. In β-$YbAlB_4$, the valence of the Yb ions fluctuates between 2+ and 3+ states on a time scale τ_f. (Bottom) Schematic showing two different scenarios for Mössbauer absorption lines depending on the relative time scales τ_0 and τ_f. If $\tau_f \gg \tau_0$ (slow valence fluctuations), two distinct x-ray frequencies are detected, leading to two distinct Mössbauer absorption lines (left), whereas if $\tau_f \ll \tau_0$ (fast valence fluctuations), a single motionally narrowed x-ray frequency will be detected in the Mössbauer absorption.

τ_f gradually decreases with increasing p, exhibiting a kink across ~p^* in between 0.5 and 1 GPa, approaching the resolution limit at 2.3 GPa (Fig. 2D). This is roughly consistent with previous $\rho(T)$ measurements in β-$YbAlB_4$ (*32*); at $T < 0.5$ K and under p, $\rho(T)$ displays $\rho \sim T^\alpha$ with $\alpha = 3/2$ below p^*, and further application of pressure increases the exponent to $\alpha = 2$, stabilizing a FL state at about 1 GPa (*32*). However, the FL temperature T_{FL} depends on p, and only for $p \sim 2.3$ GPa is the system in the FL regime at $T = 2$ K (*32*).

Slow valence fluctuations

This consistency leads us to interpret the split line shape observed in the Mössbauer spectra of the SM as unusually slow valence fluctuations

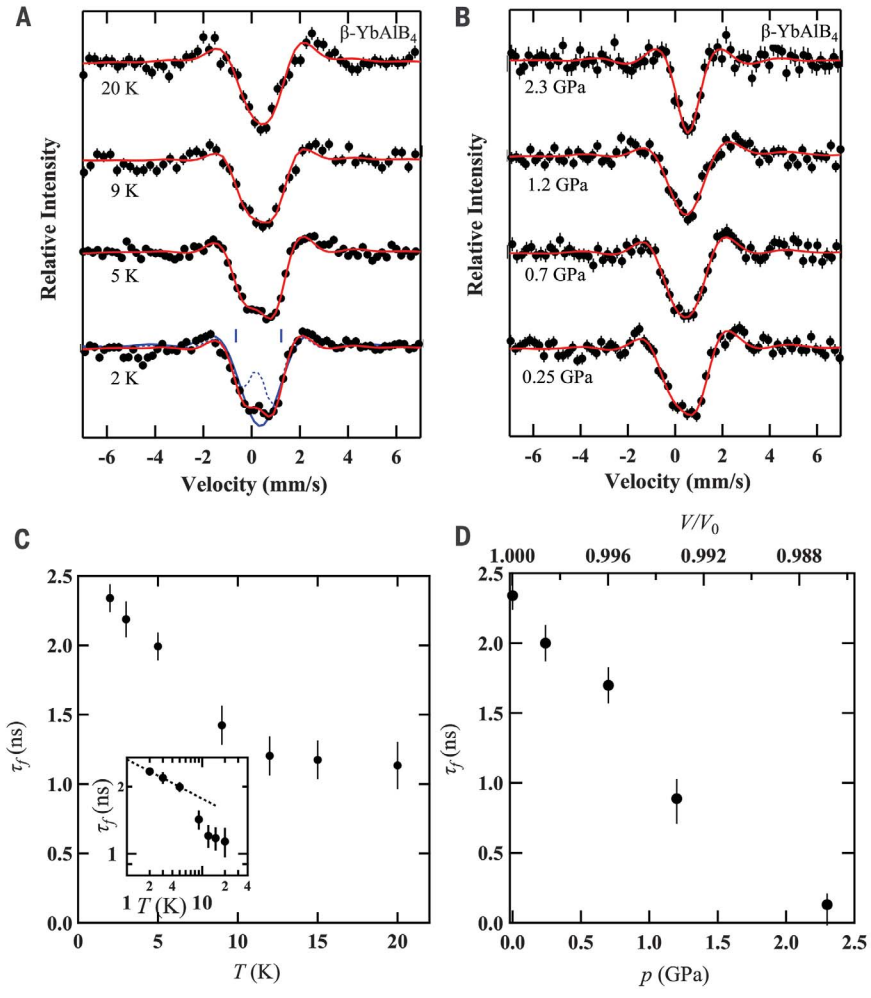


Fig. 2. Temperature and pressure dependence of synchrotron radiation–based ^{174}Yb Mössbauer spectra of β-YbAlB$_4$. (**A** and **B**) Selected spectra as a function of (A) temperature (T) at ambient pressure and (B) under external pressure (p) at 2 K. The c axis of the single-crystalline β-YbAlB$_4$ samples was aligned along the propagation vector k_0 of the incident x-ray. The solid circles with error bars and the red solid lines indicate the observed and the analytical spectra calculated as described in (*30*), respectively. In (A), the dashed blue line in the spectrum at 2 K indicates the spectrum with two static nuclear transitions expected with our experimental energy resolution, whereas the solid blue line shows a fit to the wings of the line shape, discarding the double-peak structure in the center. The deviation at the center corresponds to 5σ statistical significance (*30*). (**C**) Temperature T and (**D**) pressure p dependences of the refined fluctuation time τ_f between two different Yb charge states in β-YbAlB$_4$. (Inset) Log-log plots of τ_f versus T in β-YbAlB$_4$. The dashed line indicates $\tau_f \sim T^{-0.2}$.

between the Yb^{2+} and Yb^{3+} ionic-like states in β-YbAlB$_4$, on a time scale $\tau_f > 1$ ns that follows an approximate power-law growth $\tau_f \sim T^{-0.2}$ with decreasing temperature below T^*. The Yb^{3+} ground state is a $J_z = \pm 5/2$ moment as deduced by varying the incident angle of the x-ray (*30*). The slow charge fluctuations extend up to p^*, beyond which a conventional valence fluctuation state with rapid charge fluctuation takes over in the pressured regime corresponding to the FL regime.

The unusual aspect of the observed charge dynamics is that not only are they slower than the Planckian time $\tau_f \gg \tau_{tr} \sim 10^{-2}$ ns at 2 K, they are also slower than the characteristic time scale of the lattice vibrations. Therefore, the lattice is expected to adiabatically respond to the associated charge redistribution. Each valence fluctuation of Yb atoms is then dressed by N_p phonons, leading to the formation of a polaron (*38*, *39*) and renormalizing the matrix element for the charge fluctuations; this provides a mechanism for enhancing their time scale ($\tau_f \rightarrow \tau_f e^{N_p}$) (*30*). Analysis of the Mössbauer spectra allowed us to directly check this scenario. We used the T-dependence of the absorption components in the spectra to determine the Lamb-Mössbauer (recoil-free) factor f_{LM} in β-YbAlB$_4$, which is the equivalent of the Debye-Waller factor in a usual scattering experiment. Generally, $-\ln f_{LM} = k_0^2 \langle \Delta z^2 \rangle$, where Δz is an atomic displacement from a regular position in a crystal along the direction of k_0 (*40*). The expression for the variance in atomic position is

$$\langle \Delta z^2 \rangle \propto \int_0^\infty d\omega \frac{F(\omega)}{\omega} \overbrace{\left[\frac{1}{2} + \frac{1}{e^{\omega/T} - 1}\right]}^{\frac{1}{2}\coth(\beta\omega/2)} \quad (1)$$

where $F(\omega)$ is the (partial) phonon density of states. In a Debye model, $F(\omega) \propto \omega^2$, which leads to $\langle \Delta z^2 \rangle \propto \left[3/2 + (\pi/\Theta_D)^2 T^2\right]$ at $T \ll \Theta_D$, where Θ_D is the Debye temperature (*30*). This Deybe relation holds above T^* at ambient pressure, where τ_f (~1.15ns) is independent of T (Fig. 3A); from this, we estimated Θ_D = 95 K, corresponding to the lattice response time $\tau_L \sim h/k_B\Theta_D \sim 0.5$ ps, so that $\tau_f \gg \tau_L$. The estimated Θ_D (= 95 K) value is smaller than that (195 K) of a conventional valence fluctuation metal $YbAl_2$ (*41*). This indicates that the lattice vibrations are softer in β-YbAlB$_4$, which suggests an enhanced effective coupling between slow charge fluctuation modes and lattice vibrations.

Additionally, in the QC regime below T^*, where τ_f develops temperature-dependence, $\langle \Delta z^2 \rangle$ departs from this Debye behavior (Fig. 3A), indicating an enhancement in the quantum fluctuations, $\langle \Delta z^2 \rangle = \langle \Delta z^2 \rangle_{Debye} + \delta\langle \Delta z^2 \rangle$, of the Yb ions. The $\sqrt{\delta\langle \Delta z^2 \rangle} \sim 0.014$ Å root mean square fluctuation observed here is comparable with the quantum fluctuations of the phonon mode, which is around 0.05 Å estimated from $\frac{1}{4} k_B\Theta_D \sim \frac{1}{2} m_{Yb} (k_B\Theta_D/h)^2 \langle \Delta z^2 \rangle$. $\langle \Delta z^2 \rangle$ is approximately constant at 2 K for $p < p^*$ and then drops when $p > p^*$ (Fig. 3B), indicating that the anomalous vibrations of the lattice, $\delta\langle \Delta z^2 \rangle$, disappear in the FL regime at low temperatures.

The saturation of $\langle \Delta z^2 \rangle$ for $T < T^*$ and $p < p^*$ implies that the phonon spectrum $F(\omega)$ has changed its form to compensate the coth(βω/2) term in the integral (1). This then suggests that at energies and temperatures below T^*, $F(\omega)$ acquires a temperature-dependence $F(\omega, T) = \varphi(\omega)\tanh(\omega/2T)$ that cancels the coth(ω/2T) term in integral (1). The function tanh(ω/2T) ~ ω/2T for $\omega \ll T$, and tanh(ω/2T) ~ 1 for $\omega \gg T$, and thus has the marginal Fermi liquid (MFL) form. This enhancement in phonon density of states should be observable in inelastic neutron scattering measurements. Because the phonons are linearly coupled to the charge density of the electrons, the appearance of a MFL component in the phonon spectrum is an indication of MFL behavior in the charge fluctuations. The enhancement of τ_f through polaron formation is crucial for slowing the

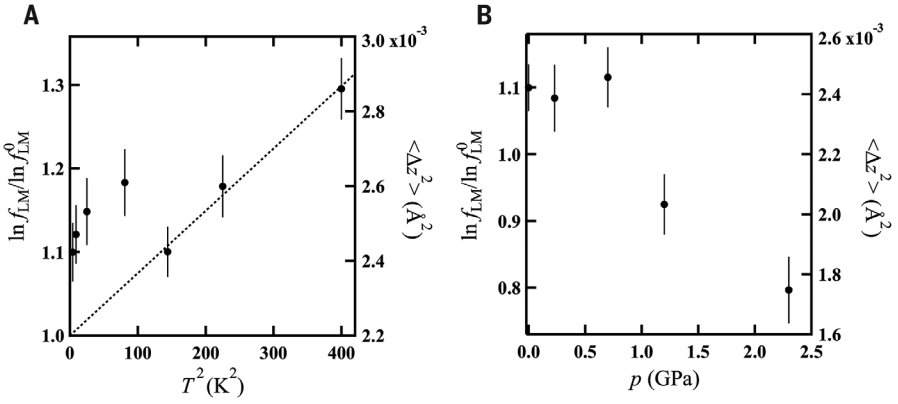


Fig. 3. Lamb-Mössbauer factor f_{LM} for β-YbAlB$_4$. (**A** and **B**) Shown is ($\ln f_{LM}/\ln f^0_{LM}$) as a function of (A) T^2 at ambient pressure and (B) under external pressure p at 2K. The Lamb Mössbauer factor f_{LM} is determined from the intensity of the Mössbauer absorption line. Its logarithm ($\ln f_{LM}/\ln f^0_{LM}$) measures the root-mean-squared fluctuations $\langle \Delta z^2 \rangle$ in the z axis position of the nucleii in units of the zero-point fluctuations (T = 0) of the Debye model. In (A) and (B), $\ln f^0_{LM}\left(\propto -\frac{3}{2}\frac{E_R}{k_B\Theta_D}\right)$ was obtained by extrapolating the high-temperature behavior to zero temperature at ambient pressure, where $E_R\left(=\frac{E_\gamma^2}{2m_{Yb}c^2}\right)$ is recoil energy. In (A), the dashed line indicates a linear relation between $\ln f_{LM}$ and T^2. For ^{174}Yb Mössbauer resonance of k_0 = 38.75Å^{-1}, $\langle \Delta z^2 \rangle$ for the Yb ions was evaluated in β-YbAlB$_4$ from the T and p dependences of $\ln f_{LM}/\ln f^0_{LM}$ by using Θ_D = 95 K. In (A) and (B), the $\langle \Delta z^2 \rangle$ values (right axis) are ~2.6 × 10^{-3} Å^2 in the SM regime. In (B), $\langle \Delta z^2 \rangle$ decreases to 1.7 × 10^{-3} Å^2 in the pressured regime corresponding to the FL regime, which is comparable with that for YbAl$_2$ (*41*).

charge fluctuations down to time scales accessible to Mössbauer spectroscopy.

Discussion and outlook

A possible interpretation of our results is the QC tuning of a critical end point of a classical valence transition (*42*) between the Yb^{2+} and Yb^{3+} ionic states. Such first-order valence transition lines, with second-order end points, are well established in rare earth compounds. It has been suggested (*42*) that the tuning of such an end point to zero temperature may provide an explanation of the observed Mössbauer spectra.

An alternative interpretation is that the observed valence fluctuation modes are an intrinsic property of the SM regime connected with a spin charge separation that develops with the collapse of the *f*-electron Fermi surface (*43*–*46*). This scenario suggests that similar slow charge fluctuations will be manifested in the Mössbauer spectra of any partial Mott localization critical point, such as in other heavy fermions and iron-based superconductors.

We provide direct evidence for unusually slow charge fluctuations in the SM regime of β-YbAlB$_4$ by using SR-based Mössbauer spectroscopy. Because their time scales are longer than that of the lattice response, we have inferred polaronic formation in the mixed valence regime (*38*, *39*). Both the slow charge fluctuation modes and the anomalous vibrations of the lattice disappear in the pressure-induced FL regime. It is natural to expect that this observed slow charge mode is connected to the linear resistivity often observed in SMs. Various theoretical approaches (*13*, *14*) have suggested that the previously unknown transport properties of SMs are linked to the universal quantum hydrodynamics of a Planckian metal. Because the local equilibrium is established at the scale of Planckian time, it is natural to regard the slow charge fluctuations detected here as a possible signature of a distinct hydrodynamic mode. This would suggest that nanosecond charge fluctuations and anomalous vibrations are not specific to β-YbAlB$_4$ but rather are universal properties of SM regimes in quantum materials.

REFERENCES AND NOTES

1. G. R. Stewart, *Rev. Mod. Phys.* **73**, 797–855 (2001).
2. T. R. Chien, Z. Z. Wang, N. P. Ong, *Phys. Rev. Lett.* **67**, 2088–2091 (1991).
3. S. Nakatsuji *et al*., *Nat. Phys.* **4**, 603–607 (2008).
4. J. G. Analytis *et al*., *Nat. Phys.* **10**, 194–197 (2014).
5. S. Paschen *et al*., *Nature* **432**, 881–885 (2004).
6. H. Shishido, R. Settai, H. Harima, Y. Ōnuki, *J. Phys. Soc. Jpn.* **74**, 1103–1106 (2005).
7. P. Gegenwart, Q. Si, F. Steglich, *Nat. Phys.* **4**, 186–197 (2008).
8. K. Kuga *et al*., *Sci. Adv.* **4**, eaao3547 (2018).
9. F. Laliberté *et al*., *Nat. Commun.* **2**, 432 (2011).
10. E. Fradkin, S. A. Kivelson, M. J. Lawler, J. P. Eisenstein, A. P. Mackenzie, *Annu. Rev. Condens. Matter Phys.* **1**, 153–178 (2010).
11. J.-H. Chu *et al*., *Science* **329**, 824–826 (2010).
12. M. J. Lawler *et al*., *Nature* **466**, 347–351 (2010).
13. J. Zaanen, Y.-W. Sun, Y. Liu, K. Schalm, *Holographic Duality in Condensed Matter Physics* (Cambridge Univ. Press, 2016).
14. S. A. Hartnoll, A. Lucas, S. Sachdev, *Holographic Quantum Matter* (MIT Press, 2018).
15. S. A. Hartnoll, A. P. Mackenzie, *Rev. Mod. Phys.* **94**, 041002 (2022).
16. P. T. Brown *et al*., *Science* **363**, 379–382 (2019).
17. L. Prochaska *et al*., *Science* **367**, 285–288 (2020).
18. S. Vig *et al*., *SciPost Phys.* **3**, 026 (2017).
19. M. Mitrano *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **115**, 5392–5396 (2018).
20. A. A. Husain *et al*., *Phys. Rev. X* **9**, 041062 (2019).
21. O. Berkooz, M. Malamud, S. Shtrikman, *Solid State Commun.* **6**, 185–188 (1968).
22. M. Takano, N. Nakanishi, Y. Takeda, S. Naka, *J. Phys. Colloq.* **40**, C2-313 (1979).
23. M. Seto *et al*., *Phys. Rev. Lett.* **102**, 217602 (2009).
24. Y. Matsumoto *et al*., *Science* **331**, 316–319 (2011).
25. M. Okawa *et al*., *Phys. Rev. Lett.* **104**, 247201 (2010).
26. C. M. Varma, *Rev. Mod. Phys.* **48**, 219–238 (1976).
27. E. V. Sampathkumaran, *Hyperfine Interact.* **27**, 183–192 (1986).
28. R. L. Cohen, M. Eibschütz, K. W. West, *Phys. Rev. Lett.* **24**, 383–386 (1970).
29. I. Nowik, *Hyperfine Interact.* **13**, 89–118 (1983).
30. Materials and methods are available as supplementary materials.
31. A. Legros *et al*., *Nat. Phys.* **15**, 142–147 (2019).
32. T. Tomita, K. Kuga, Y. Uwatoko, P. Coleman, S. Nakatsuji, *Science* **349**, 506–509 (2015).
33. J. P. Hannon, G. T. Trammell, M. Blume, D. Gibbs, *Phys. Rev. Lett.* **61**, 1245–1248 (1988).
34. Y. Sakaguchi *et al*., *J. Phys. Soc. Jpn.* **85**, 023602 (2016).
35. P. W. Anderson, *J. Phys. Soc. Jpn.* **9**, 316–339 (1954).
36. R. Kubo, *J. Phys. Soc. Jpn.* **9**, 935–944 (1954).
37. M. Blume, *Phys. Rev.* **174**, 351–358 (1968).
38. D. Sherrington, P. Riseborough, *J. Phys. Colloq.* **37**, C4-255 (1976).
39. A. C. Hewson, D. M. Newns, *J. Phys. C Solid State Phys.* **12**, 1665–1683 (1979).
40. G. T. Trammell, *Phys. Rev.* **126**, 1045–1054 (1962).
41. D. Weschenfelder *et al*., *Hyperfine Interact.* **16**, 743–745 (1983).
42. S. Watanabe, K. Miyake, *Phys. Rev. Lett.* **105**, 186403 (2010).
43. M. Oshikawa, *Phys. Rev. Lett.* **84**, 3370–3373 (2000).
44. T. Senthil, S. Sachdev, M. Vojta, *Phys. Rev. Lett.* **90**, 216403 (2003).
45. J. H. Pixley, S. Kirchner, K. Ingersent, Q. Si, *Phys. Rev. Lett.* **109**, 086403 (2012).
46. Y. Komijani, P. Coleman, *Phys. Rev. Lett.* **122**, 217001 (2019).
47. R. Masuda *et al*., *Appl. Phys. Lett.* **104**, 082411 (2014).
48. H. Kobayashi *et al*., Observation of a critical charge mode in a strange metal. Zenodo (2023); doi:10.5281/zenodo.7542767.

ACKNOWLEDGMENTS

We thank M. Takigawa for very useful discussions and F. Iga for preparation of single-crystalline YbB$_{12}$. **Funding:** The SR-based ^{174}Yb Mössbauer experiments were performed at BL09XU and BL19LXU on SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (proposals 2011A1450, 2012B1521, 2013B1393, 2015A1458, 2016A1363, 2019B1597, and 2020A1553) and RIKEN (proposals 2016110, 20170019, 20180019, and 20190025). This work is partially supported by Grants-in-Aids for Scientific Research on Innovative Areas (15H05882 and 15H05883) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; by CREST (JPMJCR18T3); Japan Science and Technology Agency; and by Grants-in-Aid for Scientific Research (15K05182, 16H02209, 16H06345, 19H00650, and 23102723) from the Japanese Society for the Promotion of Science (JSPS); the Canadian Institute for Advanced Research; the National Science Foundation grant DMR-1830707 (P.Co. and Y.Kom.) and by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences under award DE-SC0020353 (P.Ch.). We work at the Institute for Quantum Matter, an Energy Frontier Research Center funded by DOE, Office of Science, Basic Energy Sciences under award DE-SC0019331. P.Ch. and P.C. thank S. Nakatsuji and the Institute for Solid State Physics (Tokyo) for hospitality when early stages of this work were underway. P.Ch., P.Co., and Y.Kom. acknowledge the Aspen Center for Physics and NSF grant PHY-1607611 where this work was discussed and further developed. **Author contributions:** H.Ko. designed the Synchrotron Mössbauer experiment and performed it with Y.S., H.Ki., M.O., S.I., R.M.,

Y.Kob., M.S., Y.Y., and K.T. Sample synthesis and characterization were performed by K.K., S.S., and S.N. Mössbauer analysis was carried out by H.Ko. Theoretical interpretation was provided by P.Ch., P.Co., and Y.Kom.; H.Ko., S.N., P.Ch., P.Co., and Y.Kom. contributed to writing the manuscript. **Competing interests:** The authors declare no competing interests, financial or otherwise. **Data and materials availability:** All data and simulation codes presented in this paper are deposited in Zenodo (*48*). **License information:**

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abc4787

Materials and Methods
Supplementary Text
Figs. S1 to S12
References (*49*–*75*)

Submitted 28 April 2020; resubmitted 16 June 2021
Accepted 1 February 2023
10.1126/science.abc4787

WILDFIRE EMISSIONS

Record-high CO_2 emissions from boreal fires in 2021

Bo Zheng[1,2]*, Philippe Ciais[3,1,4], Frederic Chevallier[3], Hui Yang[5], Josep G. Canadell[6], Yang Chen[7], Ivar R. van der Velde[8,9], Ilse Aben[8,10], Emilio Chuvieco[11], Steven J. Davis[7,12], Merritt Deeter[13], Chaopeng Hong[1,2], Yawen Kong[14], Haiyan Li[15], Hui Li[1], Xin Lin[3], Kebin He[2,16], Qiang Zhang[14]*

Extreme wildfires are becoming more common and increasingly affecting Earth's climate. Wildfires in boreal forests have attracted much less attention than those in tropical forests, although boreal forests are one of the most extensive biomes on Earth and are experiencing the fastest warming. We used a satellite-based atmospheric inversion system to monitor fire emissions in boreal forests. Wildfires are rapidly expanding into boreal forests with emerging warmer and drier fire seasons. Boreal fires, typically accounting for 10% of global fire carbon dioxide emissions, contributed 23% (0.48 billion metric tons of carbon) in 2021, by far the highest fraction since 2000. 2021 was an abnormal year because North American and Eurasian boreal forests synchronously experienced their greatest water deficit. Increasing numbers of extreme boreal fires and stronger climate–fire feedbacks challenge climate mitigation efforts.

Fires affect climate through direct carbon dioxide (CO_2) emissions (*1*) and multiple postfire carbon source and sink pathways (*2*, *3*), which are integral components of the global carbon cycle. Globally, CO_2 emissions from fossil fuels and land-use change averaged 9 billion metric tons of carbon (Gt C) per year since 2000 (*4*), whereas fire CO_2 emissions were approximately 2 Gt C per year (*5*). Eighty percent of the carbon released by fires is taken up by vegetation in subsequent growing seasons; the remaining 20% remains in the atmosphere much longer and contributes to the build-up of atmospheric CO_2. Fire emissions from global forest ecosystems have been increasing since 2000 (*5*), as have drought-driven extreme fires that emit vast amounts of CO_2 over forests and peatlands in short periods (*6*–*8*). Such extreme fires are projected to become more frequent and widespread in response to future climate change (*9*–*11*). Tropical forest fires have attracted much attention as a result of deforestation and forest degradation threatening biodiversity and land carbon sink. Boreal forests receive much less attention, despite being one of the most extensive and important biomes on Earth (*12*) and despite warming in the Arctic region, which is happening at a much faster rate than in the rest of the planet. The monitoring and modeling of spatiotemporal variabilities of fire emissions, especially in extensive high–carbon density ecosystems such as boreal forests, is thus critical for understanding carbon–climate feedbacks, biophysical feedbacks on Earth's surface radiation budget and temperature (*13*, *14*), and the risks to climate mitigation efforts.

However, monitoring CO_2 emissions from fires accurately and in a timely way remains challenging because of the strong perturbations of combustion environments and landscape features that occur during vegetation burning. In addition, heavy aerosol load resulting from fire smoke often interferes with satellite observations of the characteristics of landscapes created by these fires [for example, burned area (*15*) and severity (*16*)] and obscures detection of small fires (*17*). Bottom-up emission models that simulate fuel load, combustion degree, and combustion efficiency reproduce average combustion conditions and emission levels well with respect to observations from field campaigns (*1*) but do not characterize extreme burning emissions accurately (*6*). The use of fire emission factors from look-up tables to indicate biome-averaged levels (*18*, *19*) could misrepresent the mixed flaming and smoldering combustion states because of the low spatiotemporal resolution of the emission factor data (*20*). Furthermore, the coarse spatial resolution of satellite sensors used for most currently available global burned area products (typically at a resolution of several hundred meters) creates large uncertainties, particularly in areas with predominantly small fires (*17*). These shortcomings in satellites, models, and datasets hamper an accurate, dynamic, and near real-time assessment of fire carbon releases, especially those from extreme fire events.

Directly measuring CO_2 emissions from fires on the basis of satellite observations is challenging on account of the long residence time of CO_2 in the atmosphere. On the other hand, satellite observations of atmospheric carbon monoxide (CO, which has a chemical lifetime of weeks to months), derived from near- and thermal-infrared radiances, provide a powerful alternate approach to tracking fire CO_2 emissions indirectly (*21*, *22*). About 36% of global surface CO emissions are from fires and have spatiotemporal patterns distinct from other sources (anthropogenic fossil fuel and biofuel use, oceanic, and biogenic processes) and from the secondary production of CO from methane and volatile organic compounds (*23*). CO plumes from fires, especially those from extreme intensive burning, can be detected by satellites and distinguished from background CO levels because of their heterogeneous emission distribution patterns and short atmospheric lifetime. For this reason, satellite-retrieved CO column concentrations have been used to track fire emissions of different pollutants through the integration of chemical transport models, atmospheric inversion approaches, and emission ratios between CO and other species determined from in situ measurements (*6*, *23*–*28*). Advances in atmospheric inversion techniques further support the development of satellite CO-based monitoring of fire emissions.

We used satellite retrievals of CO and a global atmospheric inversion system (*5*) to estimate global weekly fire CO and CO_2 emissions at a horizontal resolution of 3.75° by 1.9°. This inversion system has been developed following continuous model development efforts over the past 15 years (*5*, *20*, *23*, *29*) and is upgraded to use the latest retrievals (version 9) of CO columns from the Measurements of Pollution in the Troposphere (MOPITT) satellite instrument (*21*, *30*), fast-track anthropogenic (*31*, *32*) and fire emission (*1*) datasets as prior information, and the European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5) global reanalysis wind fields (*33*), all of which are released with a 1-week to 2-month lag. We updated the global inversion analysis to reconstruct global fire emission anomalies from 2000 to 2021 and integrate, automate, and organize these data flows to rapidly track the dynamics of future global fire emissions with low latency. This study builds on previous work that targets global forest fire emission

[1]Shenzhen Key Laboratory of Ecological Remediation and Carbon Sequestration, Institute of Environment and Ecology, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China. [2]State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex, Beijing 100084, China. [3]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France. [4]The Cyprus Institute, Nicosia 2121, Cyprus. [5]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, 07745 Jena, Germany. [6]CSIRO Environment, Canberra, ACT 2601, Australia. [7]Department of Earth System Science, University of California, Irvine, Irvine, CA 92697, USA. [8]SRON Netherlands Institute for Space Research, Utrecht, Netherlands. [9]Department of Earth Sciences, Vrije Universiteit, Amsterdam, Netherlands. [10]Department of Physics and Astronomy, Vrije Universiteit, Amsterdam, Netherlands. [11]Universidad de Alcalá, Environmental Remote Sensing Research Group, Department of Geology, Geography and the Environment, 28801 Alcalá de Henares, Spain. [12]Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, CA 92697, USA. [13]Atmospheric Chemistry Observations and Modeling Laboratory, National Center for Atmospheric Research, Boulder, CO 80307 USA. [14]Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China. [15]School of Civil and Environmental Engineering, Harbin Institute of Technology, Shenzhen 518055, China. [16]State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China.
*Corresponding author. Email: bozheng@sz.tsinghua.edu.cn (B.Z.); qiangzhang@tsinghua.edu.cn (Q.Z.)

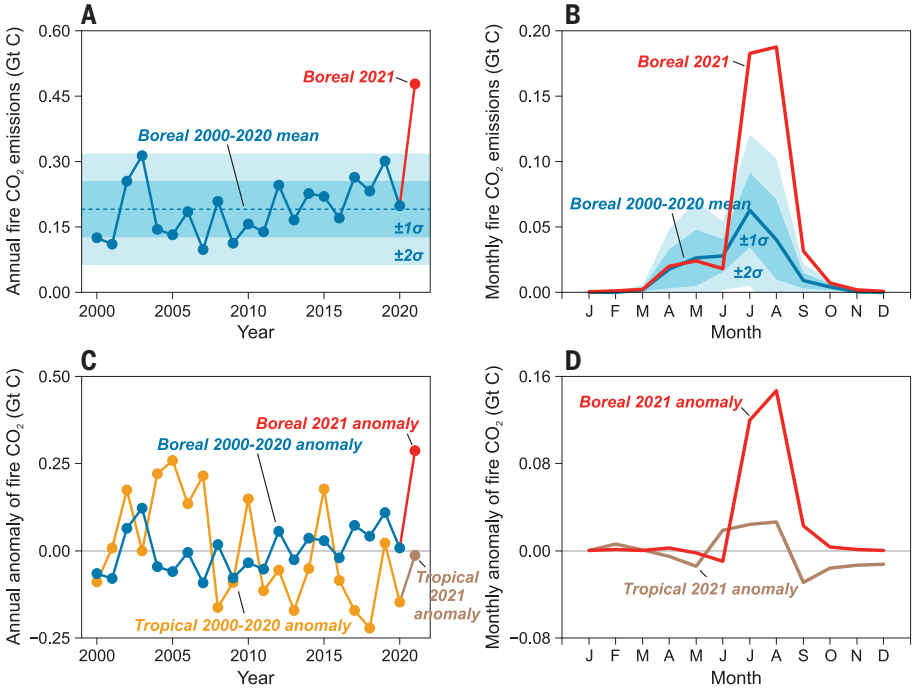


Fig. 1. Inversion estimates of boreal and tropical fire carbon emissions from 2000 to 2021. (**A** and **B**) Boreal fire annual and monthly emissions, respectively; shading indicates one and two standard deviations from the 2000 to 2020 mean. We removed the 2000 to 2020 mean to compute the annual and monthly anomalies in (**C**) and (**D**), respectively, which compare boreal emission anomalies with tropical emission anomalies, both of which are derived from our global atmospheric inversion results.

trends (*5*) and aims to dig further into regional trends and drivers on the basis of analysis enabled by the updates to our system. We focus on the trends, anomalies, and climate drivers of boreal (north of 50°N latitude) fire emissions from 2000 to 2021, given that the extensive boreal forests are vulnerable to high-latitude climate change. The global coverage of inversion results allows us to compare and understand the difference in fire emission trends between boreal and tropical ecosystems.

Extreme fire carbon release

Our inversion results highlight a significantly increasing trend of fire emissions over the boreal region at 6.6 ± 4.5 Mt C year^{-2} (nonparametric Mann-Kendall test, 95% confidence interval) from 2000 to 2021 (Fig. 1A). Even without taking 2021 into account, this increasing trend is still significant at 4.8 ± 4.2 Mt C year^{-2} since 2000 and is broadly consistent with the Global Fire Emissions Database (GFED) 4.1s estimates (*1*), which show increasing trends of 4.3 ± 3.5 Mt C year^{-2} from 2000 to 2020 and 6.3 ± 3.9 Mt C year^{-2} from 2000 to 2021. Table S1 further shows the comparable increasing fire emission trends between our inversion estimates and GFED 4.1s over both boreal North America and boreal Eurasia. The GFED 4.1s emissions after 2017 are derived from a beta data version based on the relations between NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) active fire detections and GFED 4.1s for 2003 to 2016. Our inversion results indicate that fire emissions from the tropical region (23°S to 23°N) declined from 2000 to 2021 (−9.6 ± 9.8 Mt C year^{-2}, not statistically significant) (Fig. 1C, yellow and brown curves), likely because of recently reduced savannah burning (*34*, *35*).

In 2021, boreal fires generated peak CO_2 emissions of 0.48 Gt C (0.17 Gt C from boreal North America and 0.31 Gt C from boreal Eurasia), which is 150% (equivalent to 0.29 Gt) greater than the annual mean emissions between 2000 and 2020, surpassing three standard deviations above the 2000 to 2020 average boreal fire emissions. The 2021 boreal fire emission anomaly is larger than any annual anomaly in tropical fire emissions since 2000. The GFED 4.1s database estimates show that 0.57 Gt C have been released from the boreal fires that occurred in 2021 (0.23 Gt C from boreal North America and 0.34 Gt C from boreal Eurasia), which is ~20% greater than our inversion estimates. The extremely high boreal fire emissions in 2021 occurred mainly during July and August (Fig. 1, B and D, red curves), when the monthly emissions were 0.12 and 0.15 Gt C greater than the 2000 to 2020 mean of the corresponding months, respectively. However, monthly emissions from tropical fires in 2021 show no substantial anomalies (Fig. 1D, brown curve). Our global atmospheric inversions reveal the contrasting trends between boreal and tropical fire emissions since 2000 and the record-breaking annual and seasonal anomaly of fire emissions in the boreal region superseding those of tropical regions in 2021, both of which suggest that boreal ecosystems are becoming one of the dominant source regions of intensive fires and fire carbon emissions, as well as the critical driver of positive fire–climate feedback.

The surge of boreal fire emissions in the 2021 fire season is captured by the MOPITT satellite CO observations (Fig. 2A), which show widespread large CO column anomalies (mostly 25 to 50% larger) over the entire boreal region as a result of the long-range intercontinental transport of fire CO emissions. Our atmospheric inversion system has corrected the prior fluxes to match the MOPITT CO observations. Chemical transport model simulations driven by the posterior fluxes (Fig. 2E) reproduce the CO enhancements caused by intensive fires and transported by boreal summer atmospheric circulation at high latitudes (Fig. 2B), whereas the simulations based on the climatological mean prior fluxes (2000 to 2020 mean) (Fig. 2F) largely underestimate CO column concentrations (Fig. 2C), indicating that the boreal fire emissions in 2021 are substantially higher than the 2000 to 2020 mean. We independently confirmed the occurrence of extreme fires in 2021 with 500-m resolution burned-area data (Fig. 2D) that are remotely sensed by MODIS (*36*) and with the area of forest loss due to fire in 2021, the largest since 2000 (fig. S1) (*37*). Our inversion emission results are broadly consistent with the MODIS burned area–based GFED 4.1s estimates regarding the climatological mean and trend of the boreal fire carbon emissions (fig. S2). Slight differences are shown for the 2021 emission anomaly between our inversion and GFED 4.1s, especially over boreal North America, where the large fire emission spikes in 2021 are not concentrated over the same grid cells. Such inconsistency may reflect uncertainties in the detection of burned areas and fire CO plumes by satellite, which leads to more uncertainties in bottom-up or top-down emission estimates.

Climate drivers of extreme fires

The increases in burned area and fire emissions in the 2021 boreal fire season coincide with severe heatwave, drought, and high water deficit, as shown in the ERA5 reanalysis fields (Fig. 2, G to I, and fig. S3, A to F) and the Global Fire Weather Database (fig. S3, G to I) (*38*). All regions with large fire emission spikes in 2021 experienced higher 2-m temperature anomalies (>1°C over boreal North America and >1.5°C over boreal Eurasia) (Fig. 2G). The increases in fire emissions over boreal Eurasia were also concurrent with low-precipitation anomalies (<−10 mm) (Fig. 2H) and large climatic

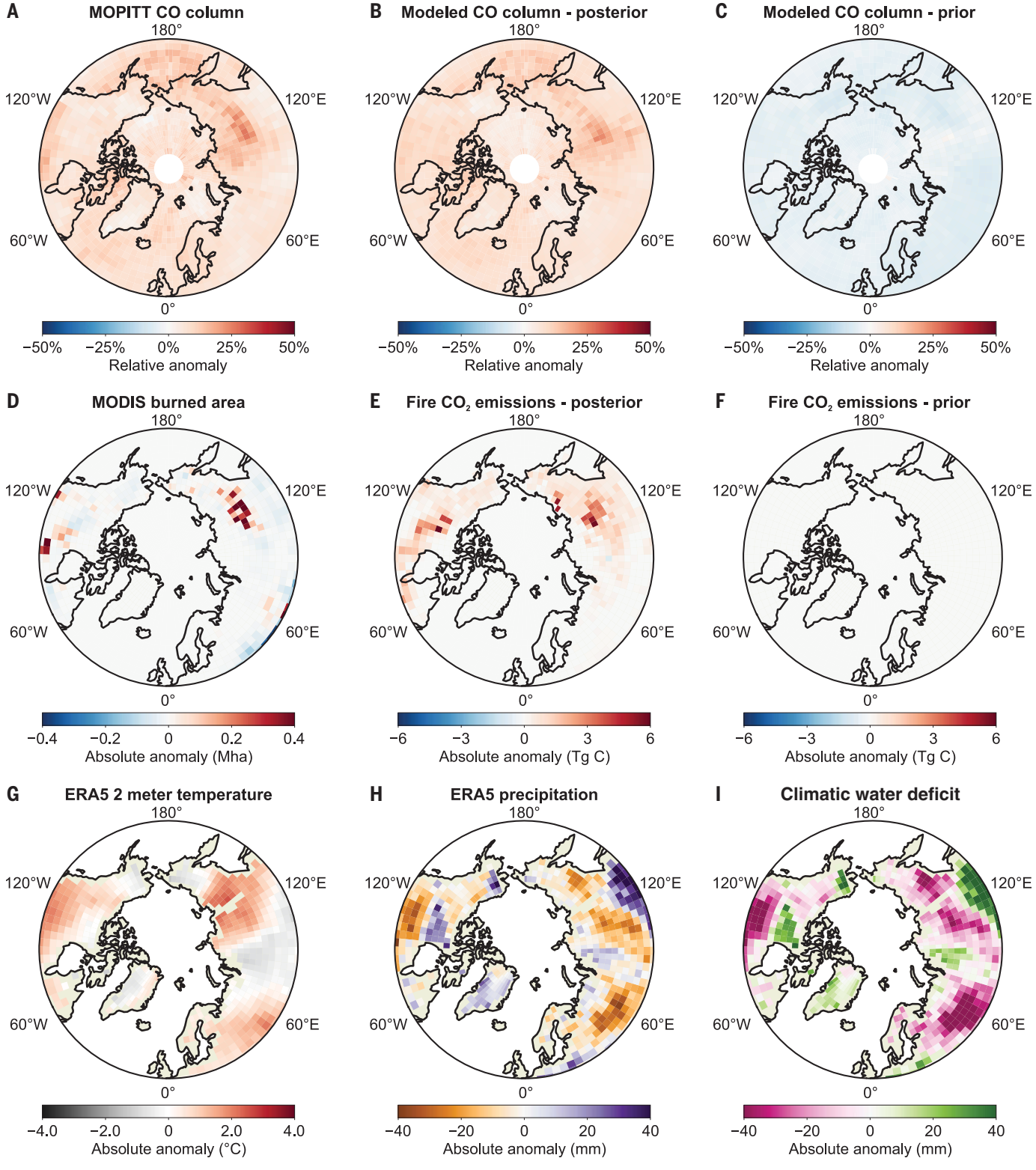


Fig. 2. The 2021 fire season anomaly (relative to the 2000 to 2020 fire season mean) of MOPITT and modeled CO columns, MODIS burned areas, inversion estimates, and weather variables in the boreal region (> 50°N). The maps show the spatial differences during the fire season (June to September) between the 2021 and 2000 to 2020 means at a horizontal resolution of 3.75° longitude by 1.9° latitude. (**A** and **D**) Satellite observations indicate MOPITT tropospheric CO column concentrations and MODIS burned areas, respectively. (**B** and **E**) Posterior estimates derived from our atmospheric inversion system, showing modeled CO columns and inversion-derived fire CO_2 emissions, respectively. (**C** and **F**) Prior data used in our atmospheric inversion system, showing modeled CO columns and fire CO_2 emissions, respectively. We use the 2000 to 2020 mean fire fluxes as prior for our inversions in 2021, so the differences shown on the map (F) are zero. (**G** to **I**) ERA5-derived monthly mean 2-m air temperature, precipitation, and climatic water deficit.

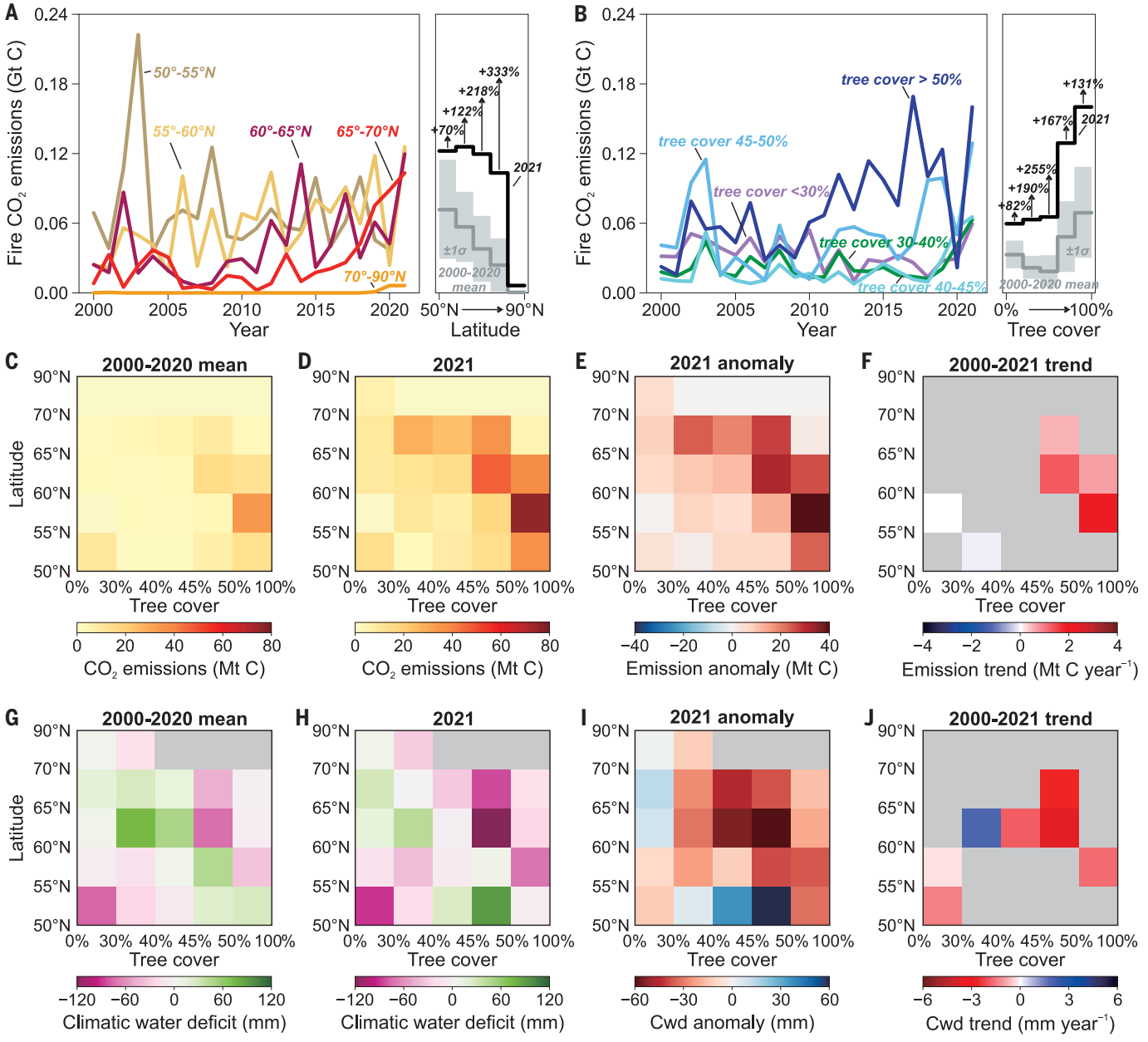


Fig. 3. Distribution of boreal fire CO_2 emissions and climatic water deficit by latitude and tree cover fraction from 2000 to 2021. (**A**) Boreal fire annual CO_2 emissions in 50° to 55°N, 55° to 60°N, 60° to 65°N, 65° to 70°N, and 70° to 90°N. (**B**) Boreal fire annual CO_2 emissions over the grid cells with tree cover fractions <30%, 30 to 40%, 40 to 45%, 45 to 50%, and >50%. (**C** to **F**) Fire emission distribution patterns and trends during the fire season (June to September) from 2000 to 2021. (**G** to **J**) Cumulative climatic water deficit during the fire season (June to September) from 2000 to 2021. Trends shown in (F) and (J) are estimated based on the nonparametric Mann-Kendall test (95% significance level); gray shadings indicate no valid data or no statistically significant trends.

water deficit anomalies (precipitation minus potential evapotranspiration, mostly <−20 mm) (Fig. 2I). Such variations in precipitation and climatic water deficit are not evident for boreal North America at the grid level (fig. S4), but we observed a sharp rise in fire emissions with climatic water deficit lower than a threshold value when investigating the regional total for boreal North America. Our data analysis implies a link between the extensive boreal fires and climate drivers (especially temperature increase or heatwaves) and also suggests that the mechanisms of wildfire propagation might be different between boreal North America and boreal Eurasia.

The geographical analysis combining latitudinal distributions of fire emissions and land cover maps suggests that the extremely high emissions in 2021 occur over the boreal regions with broad tree cover at the high northern latitudes (Fig. 3). Larger relative emission increases are observed at the higher northern latitudes (Fig. 3A) and the areas with larger tree cover fractions (Fig. 3B). The joint distribution analysis also suggests that the boreal fire emission anomaly in 2021 expands over forest zones at the 60° to 70°N range and at the 55°N to 60°N range with tree cover fractions larger than 50% (Fig. 3, C to E), which have displayed significantly increasing trends since 2000 (Fig. 3F). The increasing fire emission trends and anomalies in the boreal region occur concurrently with the emergence of warmer

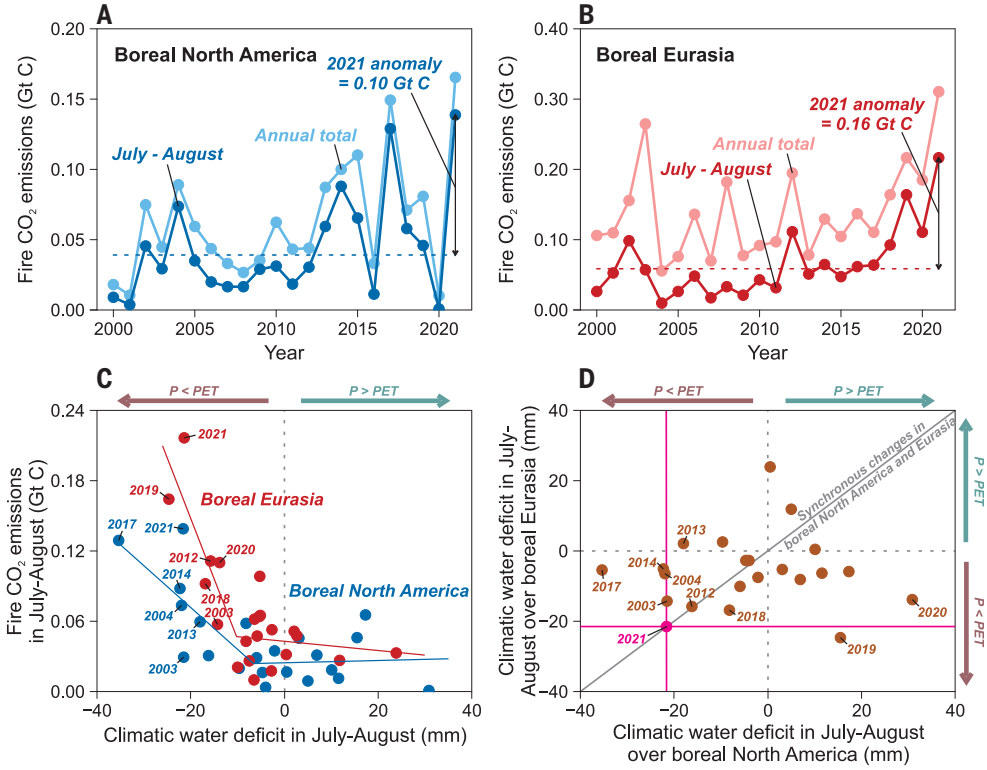


Fig. 4. Fire CO_2 emissions and climatic water deficit over boreal North America and boreal Eurasia. (**A** and **B**) Inversion-based fire CO_2 emissions of annual totals and of July to August (the 2 months with extremely high emissions in 2021) from 2000 to 2021 over boreal North America and Eurasia. Dashed lines in (A) and (B) indicate the 2000 to 2020 means of July to August fire emissions. (**C**) Relationship of fire CO_2 emissions during July to August with tree cover–weighted cumulative climatic water deficit during July to August between 2000 and 2021 for these two regions. Red and blue lines are plotted through a piecewise linear regression model; breaking points are automatically determined by means of the regression model with a least-squares algorithm. P, precipitation; PET, potential evapotranspiration. (**D**) Relationship between tree cover–weighted cumulative climatic water deficit in boreal North America (*x* axis) and boreal Eurasia (*y* axis), shown using the 2000 to 2021 data.

and drier fire seasons, which are indicated by the climatic water deficit index during fire season (Fig. 3, G to J, and fig. S5). The substantial increase in boreal fire emissions in 2021 (Fig. 3E) emerges mainly over the regions facing more severe drought conditions than those of the 2000 to 2020 average state (Fig. 3I), an increase that is more evident over boreal Eurasia. Worse, the overall trend has been moving toward drier conditions since 2000 (Fig. 3J), especially over the northern forests in the boreal region. Because high-latitude forests exhibit great vulnerability to droughts that can greatly increase fire occurrence and in turn trigger vast amounts of carbon release as in 2021 (fig. S5), the drying trends and the possible increased extreme drought frequency could drive strong positive fire–climate feedback.

Boreal North America and Eurasia

Furthermore, our study shows discrepancies in fire emission dynamics and responses to drought conditions between boreal North America and Eurasia (Fig. 4). Boreal North America shows large interannual variabilities in annual (3.2 ± 3.0 Mt C $year^{-2}$) and July to August (2.2 ± 2.4 Mt C $year^{-2}$) fire emissions (Fig. 4A), which have tended to rise since 2000 but do not show statistically significant trends, whereas boreal Eurasia shows significantly increasing annual (4.6 ± 4.1 Mt C $year^{-2}$) and July to August (3.9 ± 2.4 Mt C $year^{-2}$) emission trends since 2000 (Fig. 4B). The percentage of July to August emissions in annual total emissions, which is estimated to be flat over boreal North America (69% on average), also follows different temporal changes but has been increasing by 1.9 ± 1.0% $year^{-1}$ over boreal Eurasia since 2000. In 2021, fire emissions in July to August contribute 85 and 74%, respectively, to annual fire emissions over boreal North America and Eurasia, with these 2-month emission anomalies reaching 0.10 Gt C in North America and 0.16 Gt C in Eurasia, above the 2000–2020 means of July to August.

To quantitatively relate fire emissions to drought, we developed an indicator of tree cover–weighted climatic water deficit based on the gridded values over the two regions (Fig. 4, C and D). The indicator is designed to consider that drought over forest zones drives boreal fire emission anomalies between 2000 and 2021 (Fig. 3). Using a piecewise linear regression model built with a least-squares algorithm, we identified the breaking points at which boreal fire emissions rise sharply when the tree cover–weighted climatic water deficit falls below threshold values (slightly less than zero) (Fig. 4C). The fire carbon releases from boreal Eurasian ecosystems are more sensitive to climatic water deficit than are those from boreal North America, as illustrated by a steeper increase in fire emissions per unit of climatic water deficit decline below the thresholds. The relationship between fire emissions and climatic water deficit explains the large fire carbon releases in specific years, such as those of 2012, 2019, 2020, and 2021 over boreal Eurasia and those of 2014, 2017, and 2021 over boreal North America. 2021 was an abnormal year when boreal North America and Eurasia experienced regionally synchronous droughts with climatic water deficit lower than –20 mm in both regions, the first occurrence of this phenomenon during the studied period (Fig. 4D). North America and Eurasia do not always experience droughts simultaneously, as was the case with the drought over North America in 2017 and the drought over Eurasia in 2019, which caused extreme fires associated with large CO_2 emissions in only one of the two regions, therefore leading to smaller fire emission anomalies than those that occurred in 2021. The simple indicator we developed can help explain fire emission evolution and drivers in boreal regions and can be used to quantify and predict extreme fire emissions combined with projections of climate variables.

Climate–fire feedback

Although extreme fires have occurred frequently over forests globally in recent years, boreal fires in 2021 released an abnormally vast amount of carbon (0.29 Gt C above the 2000 to 2020 mean), more than in any previous years, according to our global inversion results (Fig. 5). This increase in fire emissions poses a widening threat to climate, given that part of the emissions might not return to vegetation and

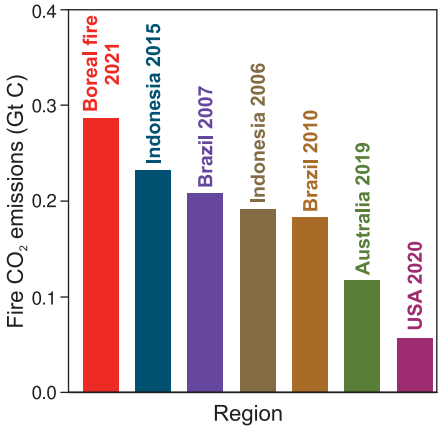


Fig. 5. Comparisons of emission anomalies from boreal fires in 2021 with those from other large extreme fire events in different years. The emission anomalies of fires shown here are calculated as the difference in emissions between the target year and the 2000 to 2020 average.

soils because of postfire regrowth failure in a warming climate (*39*). Extreme fire events are more likely to occur with global warming in the future, which could place the boreal landscape in a frequently disturbed state and substantially suppress the stable storage of carbon. This positive climate–fire feedback exacerbates incomplete postfire recovery and resequestering of carbon after fires in subsequent growing seasons.

Fires are a key dynamic of boreal regions in the Northern Hemisphere. Our study highlights a two-decade trend of expanding summer fires in boreal forests of Eurasia, producing prodigious carbon emissions and driven by persistent and worsening drought. Increases in lightning over boreal forests and tundra may also play an important role (*40*–*42*). 2021 was exceptional in terms of drought conditions, the intensity and spatial extent of fires, and associated carbon emissions, all of which reached record highs since the MOPITT satellite instrument began measurements in 2000. The hotter and drier fire seasons in boreal regions, likely due to human-driven climate change, increase the flammability of vegetation biomass and lightning ignition efficiency, promoting extreme and intense fires. These boreal forest fires release 10 to 20 times as much carbon per unit of area burned as do other ecosystems (for example, grassland), according to our global average inversion estimates, and recovery is slow. The global atmospheric CO inversion framework that we demonstrate provides fire emission estimates from a top-down perspective that could be incorporated into a fire-impacts evaluation model, which represents an important step toward an integrated system capable of monitoring and evaluating global and regional fire carbon budgets, postfire land-use fluxes, and the net impact of fire emissions on atmospheric CO_2. The satellite CO-based inversion assessment of fire carbon emissions complements and cross-validates the estimates from biogeochemical models, enhancing our confidence in the magnitude of this important component of the carbon budget.

REFERENCES AND NOTES

1. G. R. van der Werf *et al.*, *Earth Syst. Sci. Data* **9**, 697–720 (2017).
2. S. P. K. Bowring, M. W. Jones, P. Ciais, B. Guenet, S. Abiven, *Nat. Geosci.* **15**, 135–142 (2022).
3. Y. Yin *et al.*, *Nat. Commun.* **11**, 1900 (2020).
4. P. Friedlingstein *et al.*, *Earth Syst. Sci. Data* **14**, 1917–2005 (2022).
5. B. Zheng *et al.*, *Sci. Adv.* **7**, eabh2646 (2021).
6. I. R. van der Velde *et al.*, *Nature* **597**, 366–369 (2021).
7. R. D. Field *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9204–9209 (2016).
8. J. Liu *et al.*, *Science* **358**, eaam5690 (2017).
9. M. Senande-Rivera, D. Insua-Costa, G. Miguez-Macho, *Nat. Commun.* **13**, 1208 (2022).
10. J. T. Abatzoglou, A. P. Williams, R. Barbero, *Geophys. Res. Lett.* **46**, 326–336 (2019).
11. M. W. Jones *et al.*, *Rev. Geophys.* **60**, e2020RG000726 (2022).
12. M. Scheffer, M. Hirota, M. Holmgren, E. H. Van Nes, F. S. Chapin 3rd, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 21384–21389 (2012).
13. J. T. Randerson *et al.*, *Science* **314**, 1130–1132 (2006).
14. Z. Liu, A. P. Ballantyne, L. A. Cooper, *Nat. Commun.* **10**, 214 (2019).
15. E. Chuvieco *et al.*, *Remote Sens. Environ.* **225**, 45–64 (2019).
16. J. W. Kaiser *et al.*, *Biogeosciences* **9**, 527–554 (2012).
17. R. Ramo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2011160118 (2021).
18. T. T. van Leeuwen *et al.*, *Biogeosciences* **11**, 7305–7329 (2014).
19. M. O. Andreae, *Atmos. Chem. Phys.* **19**, 8523–8546 (2019).
20. B. Zheng, F. Chevallier, P. Ciais, Y. Yin, Y. Wang, *Geophys. Res. Lett.* **45**, 11,998–12,007 (2018).
21. M. N. Deeter *et al.*, *Remote Sens. Environ.* **262**, 112516 (2021).
22. R. R. Buchholz *et al.*, *Remote Sens. Environ.* **256**, 112275 (2021).
23. B. Zheng *et al.*, *Earth Syst. Sci. Data* **11**, 1411–1436 (2019).
24. Y. Yin *et al.*, *Atmos. Chem. Phys.* **15**, 13433–13451 (2015).
25. Z. Jiang *et al.*, *Atmos. Chem. Phys.* **17**, 4565–4583 (2017).
26. Y. Yin *et al.*, *Geophys. Res. Lett.* **43**, 10,472–10,479 (2016).
27. N. Nechita-Banda *et al.*, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **373**, 20170307 (2018).
28. M. Krol *et al.*, *Atmos. Chem. Phys.* **13**, 4737–4747 (2013).
29. B. Zheng *et al.*, *Environ. Res. Lett.* **13**, 044007 (2018).
30. M. Deeter *et al.*, *Atmos. Meas. Tech.* **15**, 2325–2344 (2022).
31. Z. Liu *et al.*, *Sci. Data* **7**, 392 (2020).
32. Z. Liu *et al.*, *Nat. Commun.* **11**, 5172 (2020).
33. H. Hersbach *et al.*, *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
34. N. Andela, G. R. van der Werf, *Nat. Clim. Chang.* **4**, 791–795 (2014).
35. N. Andela *et al.*, *Science* **356**, 1356–1362 (2017).
36. L. Giglio, L. Boschetti, D. P. Roy, M. L. Humber, C. O. Justice, *Remote Sens. Environ.* **217**, 72–85 (2018).
37. A. Tyukavina *et al.*, *Front. Remote Sens.* **3**, 825190 (2022).
38. R. D. Field *et al.*, *Nat. Hazards Earth Syst. Sci.* **15**, 1407–1423 (2015).
39. A. Burrell, E. Kukavskaya, R. Baxter, Q. Sun, K. Barrett, in *Ecosystem Collapse and Climate Change*, J. G. Canadell, R. B. Jackson, Eds. (Springer International Publishing, Cham, 2021), pp. 69–100.
40. Y. Chen *et al.*, *Nat. Clim. Chang.* **11**, 404–410 (2021).
41. D. L. Finney, *Nat. Clim. Chang.* **11**, 379–380 (2021).
42. T. D. Hessilt *et al.*, *Environ. Res. Lett.* **17**, 054008 (2022).
43. B. Zheng, Global fire CO_2 emissions 2000-2021, Figshare (2023); https://doi.org/10.6084/m9.figshare.21770624.

ACKNOWLEDGMENTS

Funding: This work was supported by the Young Elite Scientists Sponsorship Program by CAST grant YESS20200135 (B.Z.), National Natural Science Foundation of China grant 41921005 (Q.Z.), and Scientific Research Start-up Funds from Tsinghua Shenzhen International Graduate School grant QD2021024C (B.Z.). H.Y. was supported by grant 50EE1904 funded by the German Federal Ministry of Economics and Technology. **Author contributions:** Conceptualization: B.Z. and Q.Z. Methodology: B.Z., P.C., and F.C. Investigation: B.Z., P.C., H.Y., J.G.C., Y.C., and X.L. Visualization: B.Z. Funding acquisition: B.Z. and Q.Z. Project administration: B.Z. Supervision: B.Z., Q.Z., and K.H. Writing – original draft: B.Z. Writing – review and editing: B.Z., P.C., F.C., J.G.C., Y.C., E.C., M.D., C.H., Y.K., Ha.L., Hu.L., I.R.v.d.V., I.A., S.J.D., and Q.Z. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The MOPITT Version 9 products are available from NASA through the Earthdata portal (https://earthdata.nasa.gov/) or directly from the ASDC archive (https://asdc.larc.nasa.gov/data/MOPITT/). The GFED 4.1s fire emissions data can be derived from https://www.geo.vu.nl/~gwerf/GFED/GFED4/. The CEDS anthropogenic emissions data files can be accessed through https://data.pnnl.gov/dataset/CEDS-4-21-21. The country- and sector-level CO_2 emission growth rates are derived from the Carbon Monitor project (https://carbonmonitor.org/). The ERA5 reanalysis data can be derived from https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5. The inversion-based global gridded fire CO_2 emissions data are available from (*43*). The global atmospheric inversion system used in this study can be accessed through http://community-inversion.eu/installation.html#getting-the-code.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade0805
Materials and Methods
Figs. S1 to S5
Table S1
References (*44*–*48*)

Submitted 23 July 2022; accepted 19 January 2023
10.1126/science.ade0805

LANDSCAPE DYNAMICS

Hundred million years of landscape dynamics from catchment to global scale

Tristan Salles[1]*, Laurent Husson[2], Patrice Rey[1], Claire Mallard[1], Sabin Zahirovic[1], Beatriz Hadler Boggiani[1], Nicolas Coltice[3], Maëlis Arnould[4]

Our capability to reconstruct past landscapes and the processes that shape them underpins our understanding of paleo-Earth. We take advantage of a global-scale landscape evolution model assimilating paleoelevation and paleoclimate reconstructions over the past 100 million years. This model provides continuous quantifications of metrics critical to the understanding of the Earth system, from global physiography to sediment flux and stratigraphic architectures. We reappraise the role played by surface processes in controlling sediment delivery to the oceans and find stable sedimentation rates throughout the Cenozoic with distinct phases of sediment transfer from terrestrial to marine basins. Our simulation provides a tool for identifying inconsistencies in previous interpretations of the geological record as preserved in sedimentary strata, and in available paleoelevation and paleoclimatic reconstructions.

Landscapes are the expression of the interplay between tectono-geodynamic processes that deform and modify Earth's surface and its climate, which acts to transfer sediments from sources to sinks, through drainage networks that incise and dissect the landscape into high-frequency elevation patterns. Earth's physiography acts as a key evolving boundary condition for a plethora of geological, climatic, and biological processes and models (*1*–*4*). Traditionally, sedimentology, geochemistry, and paleontological studies have been used to decode Earth's paleoelevation. However, available observations from the geological record remain sparse, providing only a low-resolution, spatially and temporally fragmented representation of past physiography and are, to our knowledge, unable to account for past drainage networks, making it difficult to reconstruct the partitioning and distribution of sedimentary basins, geochemical cycles, the flux of nutrients from continents to oceans, and the evolution of global atmospheric and oceanic circulation. To unlock these constraints, we must be able to model—with high levels of detail—the evolution of the Earth's physiography at the global scale and through geological time.

We present a prediction of the past physiography of the Earth at a global scale, at fine resolution [10 km spatially, 1 million years (Myr) in time]. We simulate landscape evolution resulting from erosion and deposition, adjusting the physiography over time by assimilating a long-wavelength paleoelevation reconstruction derived from the geological record and paleontological archives (*5*) and paleoclimatic data from a general circulation model (*3*). Our results are first calibrated using a series of present-day observations. We then extract an ensemble of predictions that we compare with well-documented natural examples of sediments and water fluxes, stratigraphic architecture of sedimentary basins, and estimates of sediment volumes trapped in terrestrial sinks. Our numerical framework integrates the physics of surface processes and provides 100 Myr of high-spatial and temporal resolution physiography and sediment accumulation maps, fundamental in interpreting and modeling Earth's past environmental and biological conditions.

[1]School of Geosciences, The University of Sydney, Sydney, Australia. [2]CNRS, ISTerre, Université Grenoble-Alpes, Grenoble, France. [3]ENS, PSL Research University, CNRS UMR 8538, Paris, France. [4]Université Lyon 1, LGL-TPE, Villeurbanne, France.
*Corresponding author. Email: tristan.salles@sydney.edu.au

A global landscape and sediment transport geomodel

We developed a modeling framework accounting for geodynamic, tectonic, climatic, and surface processes to perform paleolandscape reconstructions at a global scale. Based on the recently released global landscape evolution model goSPL (*6*), landscape dynamics are forced with a state-of-the-art paleoelevation reconstruction tied to its plate tectonic model (*5*), and a series of paleoprecipitation maps (*3*). We implement a stepwise matching approach (see SM) to assimilate the paleoelevation dataset and estimate uplift rates over time. In this scheme, the envelope of the reconstructed landscape targets the long-wavelength pattern of the paleoelevation reconstruction model. Predictive model outputs from our paleolandscape reconstructions allow us to build high-resolution (~10 km) maps of the physiography and drainage network organization at the global scale (Fig. 1A), making it possible to simulate past sediment and fluvial paleofluxes across drainage networks (Fig. 1B), to estimate both continental and marine sediment accumulations, and to build global three-dimensional (3D) compacted stratigraphic architectures (Fig. 1C).

First, we calibrate our model (table S1) using modern estimates of suspended and bedload transport from the land to the ocean corresponding to ~20 gigatonnes (GT) per year (Fig. 2A) and contributing to >70% of the global sediment flux (*1*). These global values are similar to those of the Land2Sea database (*7*), which contain present-day estimates

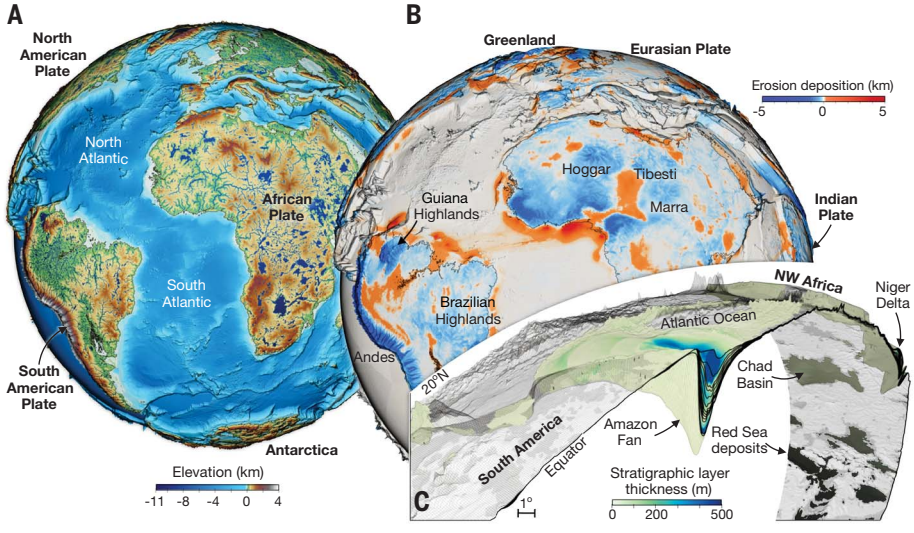


Fig. 1. Landscape evolution simulation. (**A**) Global physiography and major rivers at 50 Ma, after assimilating paleoelevation (*5*) and paleoclimatic (*3*) reconstructions. (**B**) 50 Myr of cumulative erosion and sedimentation highlighting erosion over mountain ranges (Andes) and topographic highs (Brazilian and Guiana highlands, Hoggar, Tibesti, and Marra) and major sediment accumulations along continental shelves (Atlantic margins) and endorheic basins (Amazonian watershed). (**C**) Perspective view of a stripe 20° wide featuring the relief and stratigraphy after 100 Myr of evolution; the architecture of the Amazon Fan and other sedimentary basins prominently appear.

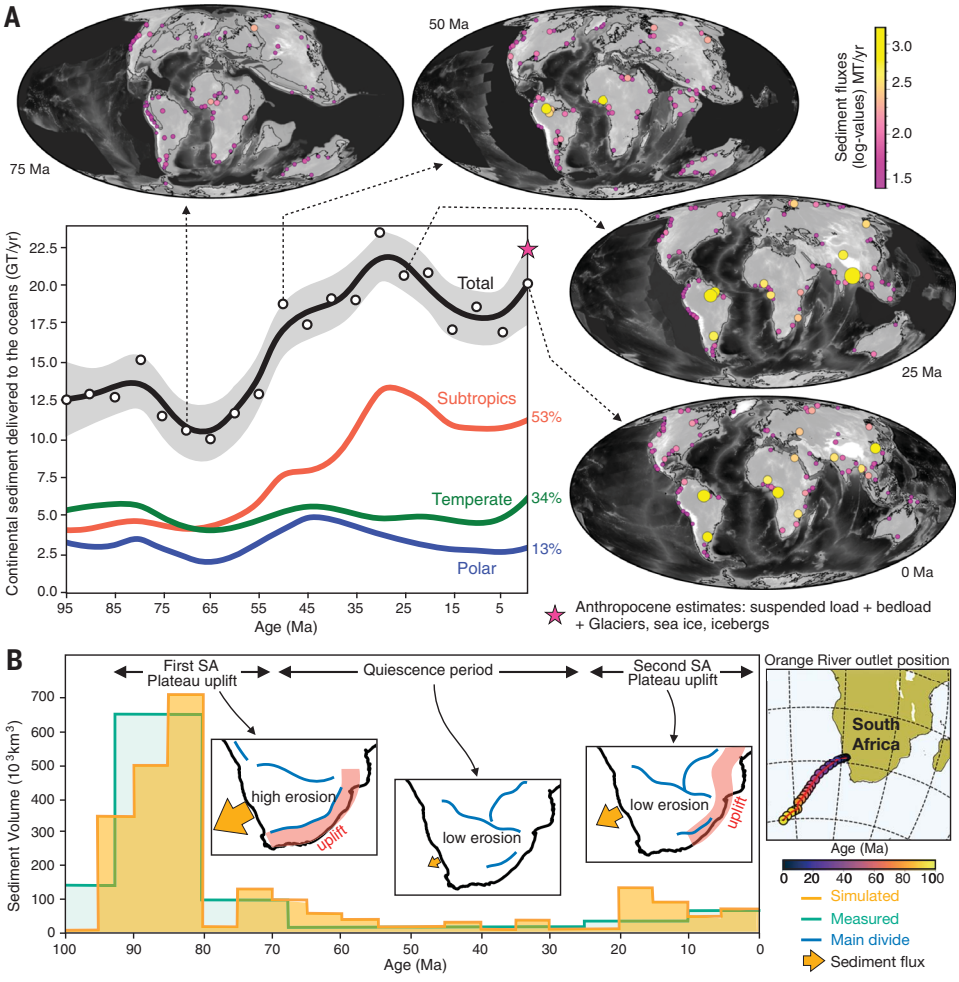


Fig. 2. Sediment flux over time. (**A**) Map view (pink and yellow circles) of the 100 highest river sediment fluxes at specific times caused by riverine and hillslope processes activated by the interplay between tectonic and climatic forcings. Latitudinal dependence of the sediment fluxes at specific times (circles) and averaged over time (Lowess regression, shaded area corresponds to confidence intervals) (subtropics: from 30°S to 30°N; temperate: 30° to 60°; and polar: above 60°). Stars show the modern estimates of sediment flux to the ocean used for model calibration in magenta (*1*) and values from the Land2Sea database in pink (*7*). (**B**) Computed (orange area) and observed (teal line) sediment fluxes in the Orange River, southern Africa over the past 100 Ma. Inset sketches show the erosion-sedimentation evolution from (*29*). Map on the right shows the Orange River mouth position reconstructed from the plate motion model (*5*).

for suspended sediment flux average to 18.6 GT per year based on 1519 exorheic river drainage basins (*7*). Latitudinal sediment flux distribution shows predominant contributions from subtropical regions corresponding to 53% of the total flux. This strong correlation between sediment flux and climate matches estimates from the global terrestrial sink catchment database (*8*) and organic carbon flux studies (*9*). From the multiple flow drainage method used to integrate runoff over the upstream drainage basin area (*6*), simulated major river positions and associated watersheds conform with those of the real world (Fig. 3A). As an example, the simulated catchment area for the Paraná Basin at the present day, after 100 Myr of simulation, has a similar shape and is only 8% smaller than the actual Paraná catchment (model resolution impedes reconstructing the highest frequencies of river sinuosity). The output also compares well with observed magnitude of discharge for the largest modern rivers (Fig. 3B) and accounts for the discharge-area scaling relationship (exponent close to 1.0 from curve fitting analysis) (*10*).

The positions of the rivers and their associated catchments are not fixed and evolve over time (Fig. 3A and fig. S6). At the catchment scale, our simulation can be used to further evaluate the transient geomorphic response of individual catchments to tectonic and climatic forcing (*11*, *12*). As an example, the main tributaries for the Paraná Basin (Grande, Paraguay, Salado, and Uruguay rivers) and their confluence positions are also well-reproduced in the simulation (Fig. 3C). From the longitudinal profiles, we identify several knickzones (Fig. 3C), the amplitudes and positions of which follow observations (*13*, *14*)—particularly for the simulated Uruguay River, which reproduces two knickzones between 1200 and 2000 km (at ~400 and 600 m elevation) similar to the present-day river profile.

Finally, our simulation also reproduces the first-order distribution of the main depocenters for the past 100 Myr (Fig. 1, B and C, Fig. 4, and fig. S10). Comparison with present-day total marine sediment thickness grid GlobSed (*15*) shows that our simulation underpredicts global thickness by 1.5 km (fig. S7, A and C), which is expected as the cumulative depositional record from this dataset spans a much longer period than ours. However, we find a better match in subtropical regions (0.8 km) where many of the largest modern oceanic sedimentary basins formed over the past 100 Myr. As an example, it predicts a maximum accumulation under the outer Pakistan shelf of up to 7.5 km in the Indus Fan (fig. S7B) that compares well with the global depocenter observed from subsurface observations (*16*). Likewise, the predicted accumulation of sediments in the post-Paleocene Bengal Fan (up to 10.3 km fig. S7B) conforms to the geological record (*17*). Similar observations can be made for the Niger Delta that accumulates in our simulation—up to 9.5 km of sediments, which can readily be compared with the 8.5-km estimate from sediment budgets (*18*) or the Eromanga Basin in central Australia with a predicted broad 1.2-km thick deposit, also within the range of observed values since the Late Cretaceous (Fig. 5B) (*19*). For the Amazon Fan (Fig. 1A), we obtain a deposit that extends up to 600 km seaward and records a 13-km thick accumulation that here again matches observations (*20*).

Water and sediment flux through space and time

The transfer of eroded sediments from the mountains to the oceans is a proxy for long-term

physical weathering and plays a central role in modulating ocean chemistry, and more generally in geochemical cycles over geological time scales (*1*, *8*, *21*, *22*). We extracted the top hundred largest sediment fluxes (Fig. 2A) and monitored water discharges over the past 100 Myr (Fig. 3A). The distribution of the water discharge, whether dominated by a few major catchments or many smaller watersheds (Fig. 3A and S6), depends on the size of the river catchments and the amount of rainfall they receive. Typically, the highest sediment fluxes correspond to the largest river discharges. However, this relationship does not always hold because the local physiography

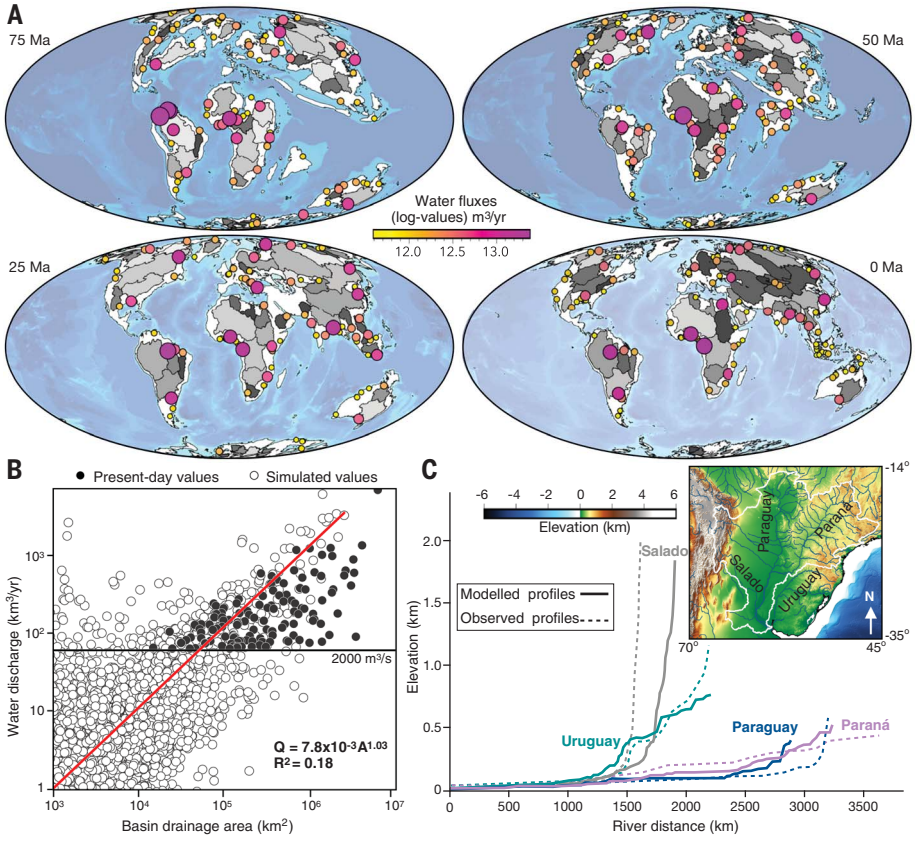


Fig. 3. Water discharges and catchment dynamics. (**A**) Map view of the 100 largest annual river discharges (circle sizes and colors scale with water flux) and 50 largest drainage basins (gray shading). (**B**) Modeled water discharge against basin drainage area (white circles) at present day (log scales). Red line, power law curve fitting. Black circles correspond to observed river fluxes at the present day for rivers whose discharge is <2000 m^3 per s (horizontal line at ~63 km^3 per yr) (*1*). Paraná drainage basin and tributaries. Simulated longitudinal main rivers profiles (solid curves) and corresponding actual river profiles (dashed, downsampled at 10 km resolution from 90 m SRTM database) (*14*). (**C**) Teal lines show identified kickzones for the Uruguay River.

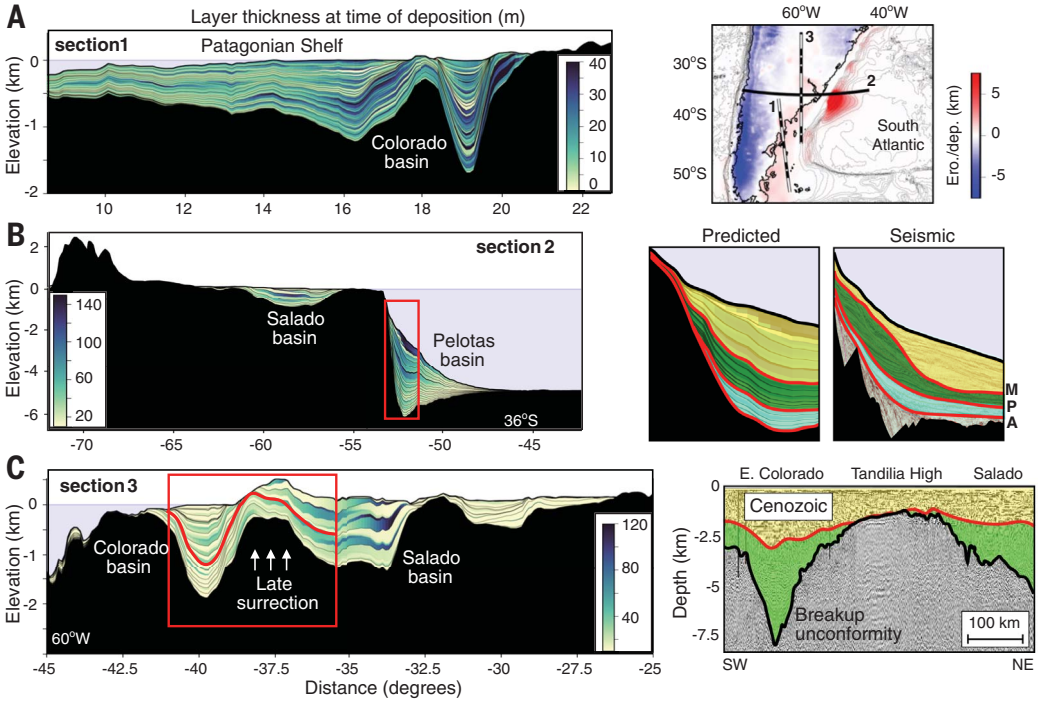


Fig. 4. Stratigraphic predictions. (**A**) Cumulative erosion/deposition (top right) at the present day, sliced with three cross sections. Simulated stratigraphy sampled every 1 Myr are displayed and colored by sediment accumulation rates (meters per million years) for each cross section. (**B**) Predicted stratigraphy for section 2 after 100 Myr of simulation (left) and focus on the Pelotas Basin (right) where comparison with megasequences from 2D seismic sections are presented (*34*). Red lines in the predicted and interpreted seismic defines the Albian (A ~100 Ma), Paleocene (P ~59 Ma), and Miocene (M ~23 Ma) transitions. (**C**) Cross section across the Salado and Colorado basins (left) compared with depth-converted interpreted seismic (*36*). Red line corresponds to the top Cretaceous horizon (~66 Ma) in both panels.

plays a crucial role in modulating sediment flux (*21*). For example, small catchments adjacent to high mountain ranges, such as on the western side of the Andes, have low water discharges but deliver a large volume of sediments to the ocean (*8*, *21*). Conversely, the northern part of South America or Central Africa at 75 Ma or Southeast Asia at 25 Ma (Fig. 3A) demonstrate that large water flux does not necessarily convey large sediment flux (Fig. 2A). This can either be attributed to high precipitation rates over relatively flat landscapes or to the storage of sediments in upstream depocenters (Fig. 5A).

Both water and sediment fluxes over time mirror several phases of continental scale drainage reorganizations. For example, our reconstruction at 75 Ma suggests that drainage networks in the upper part of the Amazon Basin were feeding an epicontinental sea at the foot of the northern Andes (Figs. 2A and 3A). In our simulation, rearrangement of these fragmented catchments into an Amazon basin that resembles the modern one occurs first around 50 Ma (fig. S9 and movie S3), followed by periods of intermittent flooding and shallow-water sedimentation in mega lakes covering the central part of the basin (movie S3). The complete drainage reversal of the Amazon River, driven by the paleoelevation reconstruction (*5*), happens during the past 20 Myr in agreement with the geological record that suggests the drying up of the Pebas system due to dynamic uplift during the Miocene epoch (*23*, *24*).

At the catchment scale we evaluate the temporal evolution of the Orange River drainage basin in southern Africa (Fig. 2B). The surface uplift history of this cratonic interior region has been attributed to mantle processes (*25*). Such regional settings, distant from plate boundaries and with negligible effects from crustal tectonics, comprise most continental areas. However, the mechanisms driving their long-term denudation patterns as well as their contributions to global sediment flux are unclear (*26*, *27*). The modeled sediment flux shows good agreement with observations (*28*, *29*); specifically, the simulation reproduces the magnitude and timing of phases of high sedimentation rates during the Late Cretaceous (93.5 to 81 Ma) and from the Oligocene to the present day (*29*).

The simulation predicts an almost twofold increase in bulk detrital flux to the oceans during the Cenozoic (Fig. 2A). The main sediment contribution comes from subtropical rivers with a threefold increase from 5 to 15 GT per year between 60 and 30 Ma. This increase is associated with the rise of the Himalayas and Tibetan plateau and to the early rearrangement of the Amazon paleobasin. Similar conclusions have been drawn when interpreting the general increase in seawater strontium isotopic composition over the past 100 Myr, with the Late Cenozoic rise of the Himalayas as its paroxysmal expression (*30*). Our results also show an increase in sediment flux delivered to the oceans over the past 15 Myr mainly from rivers in temperate regions (Fig. 2A). This increase has already been documented (*22*, *27*) but its origin remains unclear as the cooling climate during the Late Cenozoic should result in decreasing erosion rates as wetter and warmer climates accelerate mechanical rock weathering (*31*). Under these conditions, enhanced Pleistocene glacial erosion has been invoked as a major driver (*27*, *32*). Our model does not account for glacial erosion and therefore the simulated increase represents the lower range of sediment flux estimates. It shows a net positive contribution from riverine processes during the past 10 Myr (~2 GT per year representing a ~10% increase in incoming flux) that must be considered when assessing global long-term erosion rates. However, this late increase is relatively limited and the average flux over the Neogene remains relatively stable (~18.5 GT per year).

Our method offers an independent alternative to existing approaches evaluating sediment flux and long-term trends in erosion rates (*22*, *27*) such as Be isotope ratios or thermochronometric data. Similar to these approaches, our results suggest that sediment flux were higher than what they seem from the preserved sediment record (likely induced by observational biases, also known as the Sadler effect) and provide limited evidence for any major changes in worldwide erosion rates over the Late Cenozoic (*26*, *27*).

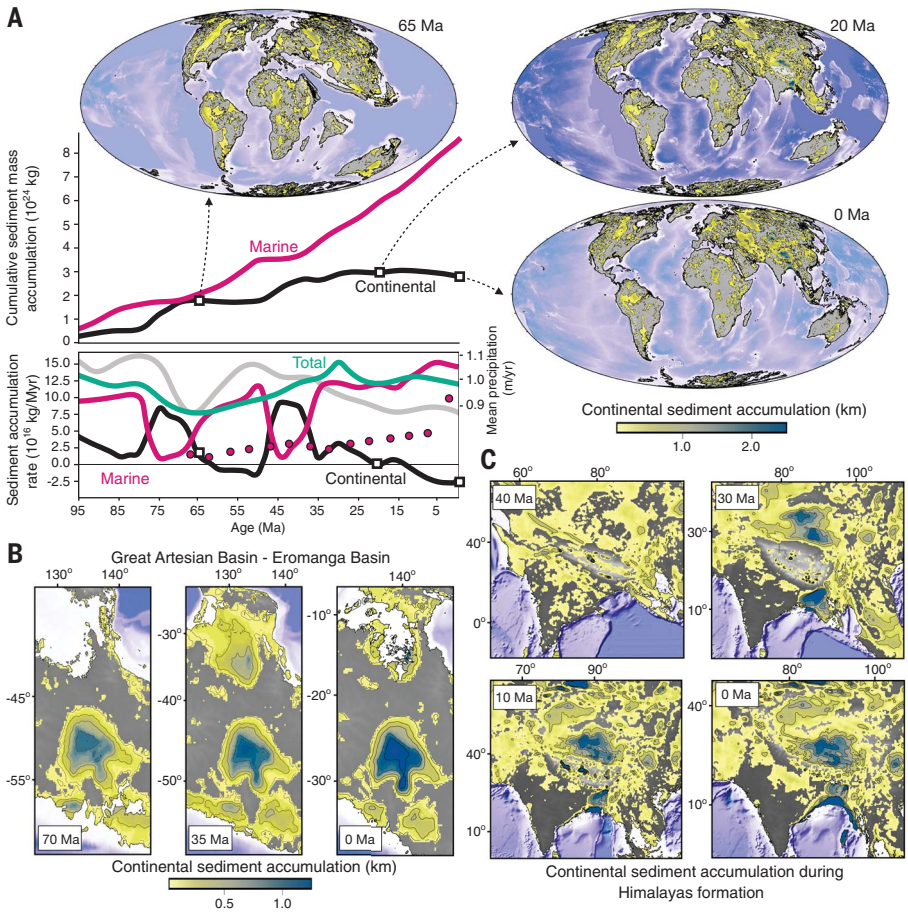


Fig. 5. Continental sediment cover temporal evolution. (**A**) Predicted temporal evolution of marine (magenta line) and continental (black line) cumulative sediment mass accumulation (top) and rate (bottom) assuming an average grain density of 2.7 g/cm^3 (total rate in teal). Magenta circles indicate estimated global values for ocean basins (*26*). The gray curve shows the average precipitation from the paleoclimate reconstruction (*3*). Maps show distribution of continental depocenters at 65, 20, and 0 Ma. (**B**) Predicted sediment accumulation in the Eromanga Basin (central Australia) showing up to 1.5 km of sediments underlying the endorheic Lake Eyre basin. (**C**) Predicted sediment accumulation in, during, and after the India-Eurasia collision. Note the formation of the Himalayan foreland basin in the south (between the Himalayan belt and the Indian continental shield) and the filling of the Tarim and Qaidam basins along the northern margin of the Tibetan plateau.

Oceanic and continental sedimentary basins

Our framework captures, at first order, the details of sedimentary basin architectures. Here, we focus again on South America (Fig. 4) and compare cross sections of the predicted stratigraphy and sediment accumulation maps with post-rift sediment thicknesses imaged by seismic data. Section 1 crosses the Colorado Basin on its northern side where the model records up to 1.8 km of sediment accumulation followed by continuous and relatively thin layers extending across the Patagonian Shelf (Fig. 4A). The simulated stratigraphic architecture of the shelf agrees with large-scale elongated stratigraphy imaged by seismic data (*33*). For section 2, the main deposits are characterized by a seaward thickening sedimentary wedge, with two depocenters that match the location, thickness, and stratigraphic record of the Salado and Pelotas basins (*34*, *35*) (Fig. 4B). Section 3 cuts through the large alluvial plain of the Paraná Basin on its northern side, where it predicts ~250 m of post-rift sediment accumulation, before entering first the Salado then the Colorado basins. In addition, our simulation predicts an early connection (up to 45 Ma) between these basins and a late tectonic forcing that uplifted the southern side of the Salado Basin by 500 m. This, however, is partly at odds with the current understanding of the tectonic and sedimentary evolution of the region (*34*, *36*). This mismatch corroborates identified inconsistencies in Late Cretaceous paleoelevation reconstructions when compared with fossil collections (*37*, *38*). It illustrates how quantitative metrics extracted from our approach could be used to test existing paleogeographic reconstructions. To adjust the paleoelevation reconstruction, one option consists of extending eastward the orographic belt of the Tandilia System during the early stages of the simulation until the predicted stratigraphic record matches the subsurface archive (*37*).

Continental sediments modulate the geological carbon cycle through weathering (*39*). As such, investigating the spatial and temporal evolution of terrestrial sinks is crucial to refining our understanding of Earth's paleoclimate variability. In contrast to marine deposits on continental margins, for which overall net deposition and good stratigraphic control exist (*15*), terrestrial sedimentary basins are typically less complete due to sedimentary erosion and hiatuses and are much harder to correlate between basins (*8*). From our simulation, we estimate the temporal evolution of endorheic (internally drained) and exorheic (coastal) continental sinks (Fig. 5A) and find that terrestrial deposits store some 28% of the total sedimentary yield over the past 100 Myr. We predict a two-stage increase, first between 80 and 60 Ma during the drying of the North America interior seaway (*5*), partial sedimentary filling of Andean retroarc foreland basins (*40*), and to a lesser extent, the transition from marine to fluvial lacustrine environments of the epicontinental Eromanga Sea (*19*) (Fig. 5B). The second, between 50 and 25 Ma, is triggered by the Himalayan orogeny and the development of extensive foreland basins (*16*, *17*) and the filling of large endorheic basins north of the Tibetan plateau (Fig. 5C). We find that global sediment accumulation rates are relatively constant over the past 100 Myr with an average value of 11.5×10^{16} kg per Myr (ranging between 7.6 and 15.0×10^{16} kg per Myr, Fig. 5A). It supports the null hypothesis (*27*) that suggests a limited impact of global climate change and mountain building events on the long-term global sedimentation rate (*41*). Despite its overall global stability, the analysis of individual contribution from either marine or continental regions tells a more complex history related to the fine balance between physiography development, sediment transport, and deposition (Fig. 5A). Over the Late Cenozoic (from 30 Ma), we predict a long-term decrease in preserved continental sediment volume. Negative rates in Fig. 5A represent erosive periods (particularly along continental margins) that translate into stratigraphic hiatuses and explain the observed mismatch between estimated and simulated sediment accumulation rates (Fig. 5A). We first relate this long-term continental decrease to limited accommodation space in major endorheic basins due to rapid infilling by sustained erosion before the Oligocene-Eocene transition. This infill is concomitant with the development of major continental-scale drainage systems throughout the Miocene (such as the Amazon, Ganges-Brahmaputra, or Indus river basins Figs. 3A, 5C, and fig. S7B) that wash out sediments from the surface of continents. In addition, the Oligocene-Eocene transition marks the emergence of continental shelves which become prone to continental erosion and increase sediment transfer to the marine environment (*41*). Our results not only highlight the critical role played by the sediment routing systems in modulating long-term sediment fluxes at the global scale but also offer a new perspective on the apparent discrepancy between Late Cenozoic stability of global erosion rates (*27*) and the observed increase in marine sedimentation (*22*, *30*).

Conclusion

We have built a quantitative and integrated exploratory modeling framework that provides global predictions of physiographic changes at 10 km and 1 Myr resolution. This time-lapse view of the Earth's landscapes is tightly linked to its associated global stratigraphy for the past 100 Myr. Our simulations yield suites of continuous, high-resolution global elevation and erosion/deposition accumulation maps, as well as water and sediment fluxes and stratigraphic information that form the basis for estimating the rates at which the global landscape changes. This dataset permits testing of the responses of the Earth's surface to different hypotheses related to climate, tectonics, and paleogeography. In addition, our model reproduces the key components of source-to-sink systems from catchment dynamics depicting river networks over time, to marine and continental sedimentary basin evolution under various forcing conditions. This analysis permits better quantification of the role that the constantly evolving physiography of the Earth has played in modulating the transport of sediments from mountain tops to ocean basins, ultimately regulating the carbon cycle and Earth's climate fluctuations through deep time. Exploring these results in tandem with the geological record will permit testing of long-standing hypotheses regarding first order features of the Earth system, such as biogeochemical cycles or biological evolution. The approach can also help to identify inconsistencies in our understanding of the origin and evolution of particular sedimentary strata and to pose alternative hypothesis that could be tested with, for example, new thermochronological or stratigraphic data in certain model-highlighted locations.

REFERENCES AND NOTES

1. J. P. Syvitski, S. D. Peckham, R. Hilberman, T. Mulder, *Sediment. Geol.* **162**, 5–24 (2003).
2. E. Straume, C. Gaina, S. Medvedev, K. Nisancioglu, *Gondwana Res.* **86**, 126–143 (2020).
3. P. Valdes, C. Scotese, D. Lunt, *Clim. Past* **17**, 1483–1506 (2021).
4. O. Hagen *et al.*, *J. Biogeogr.* **46**, 1792–1807 (2019).
5. C. Scotese, N. Wright, PALEOMAP Paleodigital Elevation Models for the Phaerozoic (2018); https://www.earthbyte.org/webdav/ftp/Data_Collections/Scotese_Wright_2018_PaleoDEM/Scotese_Wright2018_PALEOMAP_PaleoDEMs.pdf.
6. T. Salles, C. Mallard, S. Zahirovic, *J. Open Source Softw.* **5**, 2804 (2020).
7. B. Peucker-Ehrenbrink, *Geochem. Geophys. Geosyst.* **10**, 6 (2009).
8. B. Nyberg, R. Gawthorpe, W. Helland-Hansen, *Geomorphology* **316**, 1–23 (2018).
9. W. Ludwig, J.-L. Probst, S. Kempe, *Global Biogeochem. Cycles* **10**, 23–41 (1996).
10. K. Whipple, G. Tucker, *J. Geophys. Res.* **104**, 17661–17674 (1999).
11. J. Perron, L. Royden, *Earth Surf. Process. Landf.* **38**, 570–576 (2013).
12. B. A. Adams, K. X. Whipple, A. M. Forte, A. M. Heimsath, K. V. Hodges, *Sci. Adv.* **6**, eaaz3166 (2020).
13. A. G. Lima, A. L. Binda, *J. S. Am. Earth Sci.* **48**, 262–270 (2013).
14. V. Rodríguez Tribaldos, N. J. White, G. G. Roberts, M. J. Hoggard, *Geochem. Geophys. Geosyst.* **18**, 2321–2353 (2017).
15. E. O. Straume *et al.*, *Geochem. Geophys. Geosyst.* **20**, 1756–1772 (2019).
16. P. Clift, *Prog. Earth Planet. Sci.* **4**, 39 (2017).
17. M. Blum *et al.*, *Sci. Rep.* **8**, 7973 (2018).
18. J.-L. Grimaud, D. Rouby, D. Chardon, A. Beauvais, *Basin Res.* **30**, 169–186 (2018).
19. C. Braz *et al.*, *Basin Res.* **33**, 3378–3405 (2021).
20. M. Rodger, A. Watts, C. Greenroyd, C. Peirce, R. Hobbs, *Geology* **34**, 1081–1084 (2006).
21. J. D. Milliman, J. P. M. Syvitski, *J. Geol.* **100**, 525–544 (1992).
22. P. Molnar, *Annu. Rev. Earth Planet. Sci.* **32**, 67–89 (2004).

23. G. Shephard, L. Liu, D. Müller, M. Gurnis, *Gondwana Res.* **22**, 658–663 (2012).
24. N. Flament, M. Gurnis, D. Müller, D. Bower, L. Husson, *Earth Planet. Sci. Lett.* **430**, 9–18 (2015).
25. J. R. Stanley, R. M. Flowers, *Lithosphere* **12**, 74–87 (2020).
26. J. K. Willenbring, F. von Blanckenburg, *Nature* **465**, 211–214 (2010).
27. J. Willenbring, D. Jerolmack, *Terra Nova* **28**, 11–18 (2015).
28. J. Stanley *et al.*, *Journal of Geophysical Research: Solid Earth* **126**, e2020JB021243 (2021).
29. G. Baby, F. Guillocheau, J. Braun, C. Robin, M. Dall'Asta, *Terra Nova* **32**, 53–65 (2020).
30. J. M. Edmond, *Science* **258**, 1594–1597 (1992).
31. M. C. Eppes *et al.*, *Geophysical Research Letters*, **47**, 2020GL089062, (2020).
32. M. Koppes *et al.*, *Nature* **526**, 100–103 (2015).
33. J. Cavallotto, R. Violante, F. Hernández-Molina, *Biol. J. Linn. Soc.* **103**, 346–362 (2011).
34. B. Conti *et al.*, *Mar. Pet. Geol.* **83**, 1–25 (2017).
35. U. Schattner, F. J. Lobo, A. López-Ouirós, J. L. Nascimento, M. Mahiques, *Basin Res.* **32**, 293–301 (2019).
36. J. Lovecchio *et al.*, *Terra Nova* **30**, 359–368 (2018).
37. M. J. Arrouy *et al.*, *Sci. Rep.* **6**, 30590 (2016).
38. W. Cao *et al.*, *Biogeosciences* **14**, 5425–5439 (2017).
39. J. Krissansen-Totton, D. C. Catling, *Nat. Commun.* **8**, 15423 (2017).
40. M. C. Menegazzo, O. Catuneanu, H. K. Chang, *Mar. Pet. Geol.* **73**, 131–156 (2016).
41. R. Schumer, D. Jerolmack, *J. Geophys. Res.* **114**, F00A06 (2009).

ACKNOWLEDGMENTS

This research was undertaken with resources from the National Computational Infrastructure supported by the Australian Government and from Artemis HPC supported by the University of Sydney. We would also like to thank the two anonymous reviewers for their suggestions and comments. **Funding:** This work was supported by the following: Australian Research Council grant IC190100031 (to T.S.) and Australian Research Council grant DE210100084 (to S.Z.) **Author contributions:** Conceptualization: T.S. and L.H. Methodology: T.S. and L.H. Investigation: T.S. and L.H. Visualization: T.S. and L.H. Funding acquisition: T.S. and S.Z. Project administration: T.S. Supervision: T.S. Writing – original draft: T.S., L.H., P.R., N.C., and M.A. Writing – review and editing: T.S., L.H., P.R., N.C., M.A., S.Z., C.M., and B.H.B. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** The scientific software used in this study, goSPL (*13*), is available from https://github.com/Geodels/gospl and the software documentation can be found at https://gospl.readthedocs.io. We also provide a series of Jupyter notebooks used for pre-and post-processing model outputs and can be accessed at https://github.com/Geodels/gospl-global-workflows. The PALEOMAP paleoelevation reconstruction and related GPlates rotation and geometry files are downloaded from https://www.earthbyte.org/paleodem-resource-scotese-and-wright-2018. Paleoprecipitation maps from the HadCM3BL-M2.1aD model (*3*) are available at https://www.paleo.bristol.ac.uk. Simulation results provided as netcdf files containing four parameters: elevation, filled flow accumulation (rivers with lakes), flow accumulation (rivers only) and watershed (drainage basins) are available at 5 Myr interval from https://osf.io/5aek4/.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add2541
Materials and Methods
Figs. S1 to S11
Table S1
Movies S1 to S3
References (*42*–*72*)

Submitted 2 June 2022; accepted 12 January 2023
10.1126/science.add2541

ACTIVE MATTER

Lane nucleation in complex active flows

Karol A. Bacik[1]*, Bogdan S. Bacik[2], Tim Rogers[1]*

Laning is a paradigmatic example of spontaneous organization in active two-component flows that has been observed in diverse contexts, including pedestrian traffic, driven colloids, complex plasmas, and molecular transport. We introduce a kinetic theory that elucidates the physical origins of laning and quantifies the propensity for lane nucleation in a given physical system. Our theory is valid in the low-density regime, and it makes different predictions about situations in which lanes may form that are not parallel with the direction of flow. We report on experiments with human crowds that verify two notable consequences of this phenomenon: tilting lanes under broken chiral symmetry and lane nucleation along elliptic, parabolic, and hyperbolic curves in the presence of sources or sinks.

When two groups of pedestrians move past each other in opposite directions, the crowd can spontaneously segregate into contraflowing lanes (*1*–*9*). Moving in lanes reduces the risk of a collision and increases the efficiency of motion, but lane formation does not require conscious optimization effort. Indeed, the spontaneous emergence of lanes is also observed for driven colloids (*10*, *11*) and complex plasmas (*12*, *13*), and this phenomenon is hypothesized as a key to facilitating bidirectional intercellular transport in elongated domains such as axons (*14*). Provided the crowd density is low enough to avoid jamming, lane formation is also robustly reproduced by cellular automata (*15*, *16*), as well as both lattice (*17*, *18*) and off-lattice (*19*–*28*) agent-based simulations. The ubiquity of lane formation across a broad class of physical systems and the fact that the emergence of lanes in numerical simulations is largely independent of their microscopic implementation suggest the existence of a universal mechanism that underpins lane formation. Nevertheless, notwithstanding substantial progress over many years, the scientific community has not yet reached consensus about the physical origin of lane formation (*19*, *22*–*24*, *29*).

The spontaneous emergence of an apparently regular pattern suggests that a continuum field-theory approach to crowd modeling may be fruitful, through which we might expect to observe a hydrodynamic instability around the homogeneous mixed state. Calculations have previously been performed using ad hoc partial differential equation models based on heuristic reasoning, which identified at least three different plausible mechanisms of lane nucleation: gradient-induced drift (*19*, *22*, *29*), undulation-induced diffusion (*11*, *23*), and density-induced drift when chiral symmetry is broken (*30*). To settle this debate, we need to make a direct connection between the microscopic interaction rules of a given system and the emergent partial differential equations at the macroscale. Despite some notable advances for one-component systems (*31*, *32*), for active two-component flows, this has only been performed for the repulsive torque model (*33*) and also in the case of extremely soft and dense particles (*24*), which do not exhibit a laning instability.

We introduce a theoretical approach that uses temporal coarse-graining akin to Einstein's kinetic theory of Brownian motion (*34*). Our averaging scheme is valid in the case of nonjamming mixtures of hard particles, where the dynamics is dominated by pairwise interactions, which is a good approximation for typical pedestrian flows as well as dilute colloids. We recover and unify in a systematic manner the fundamental insights of Helbing and Vicsek (*22*) as well as Vissers *et al.* (*11*) and Klymko *et al.* (*23*) by showing that undulation-induced drift and diffusion can both contribute to lane nucleation. We also demonstrate that diffusive processes suppress the formation of very narrow lanes, thereby providing a dynamical selection mechanism that favors the nucleation of lanes of a particular width. We provide explicit formulas for the propensity of a given system to nucleate lanes, and we present a simple approximate rule that lanes emerge at a rate proportional to the product of agent speed, density, and an effective parameter related to the average magnitude of lateral displacement in agent-agent collisions.

The transport equations for agent densities that we derive are also readily extended to the case of systems that exhibit broken chiral symmetry, for example, pedestrians with a preference for turning right (or left) when dodging each other (*30*, *35*, *36*), rotating robots (*37*), or nonspherical driven particles. We identify an additional density-induced mechanism [previously suggested in, for example, (*30*)] that is the dominant driver of laning in the case of strongly asymmetric interactions and, in even mildly chiral systems, is responsible for producing lanes that are no longer parallel to the direction of motion. This interesting prediction is confirmed in an experiment with human crowds.

Finally, we show that lanes that are not parallel to the direction of motion also occur in situations when the trajectories of the two agent types cross at an angle. This observation encompasses a range of more complex scenarios with point targets and sources, which are particularly germane to pedestrian flows—for example, entrances and exits into open spaces. We demonstrate both theoretically and experimentally that, in this case, the laning instability is triggered along curves describing conic sections with different experimental scenarios that achieve parabolic, elliptic, or hyperbolic lane nucleation.

Our theory is designed to give insight into the fundamental mechanisms of lane formation in dilute two-component flows. Although we have tested some predictions of the theory in experiments with human crowds, it should be emphasized that our model does not incorporate a range of system-specific details that may have important effects in different settings. These include hydrodynamic effects in colloidal suspensions (*25*) or complex stimulus-response mechanisms (*38*), anticipation (*39*), and gait mechanics (*40*) in pedestrians. We also do not address multibody interactions that arise in high-density regimes, so effects such as jamming (*23*, *41*, *42*) [discussed in more detail in section IV of (*43*)] are not captured, and we have not examined the role of boundaries, which are important in many pedestrian scenarios (*3*, *44*). Such features could be incorporated into more specialized versions of our model with a specific setting in mind, but here we present only the simplest form in order to elucidate the underlying mathematics and show that many salient features of lane nucleation can be successfully explained by a simple and interpretable theory.

Kinetic description

Consider two groups of agents (labeled "+" and "−") moving in opposite directions at equal speeds in a bounded two-dimensional domain. An agent's intended motion across the domain is interrupted by encounters with agents of the opposite type; we seek to compute the statistics of many such events accumulated over a period of time Δt. Tracking a focal agent i of type "+", we obtain the approximate expression

$$\Delta \boldsymbol{r}_i^+ = \boldsymbol{v}^+ \Delta t + \sum\nolimits_{\text{collisions}} \boldsymbol{G}\left(\boldsymbol{r}_i^+ - \boldsymbol{r}_j^-\right) \quad (1)$$

where $\boldsymbol{r}$ is the position vector, $\boldsymbol{v}$ is the preferred velocity, and $\boldsymbol{G}$ is the so-called collisional operator, which captures the effects of interactions. The first term on the right-hand side describes the intended motion of an agent in

[1]Centre for Networks and Collective Behaviour, Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK. [2]Department of Human Motor Behavior, Institute of Sport Sciences, Academy of Physical Education in Katowice, ul. Mikołowska 72c, 40-065 Katowice, Poland.
*Corresponding author. Email: t.c.rogers@bath.ac.uk (T.R.); kab77@bath.ac.uk (K.A.B.)

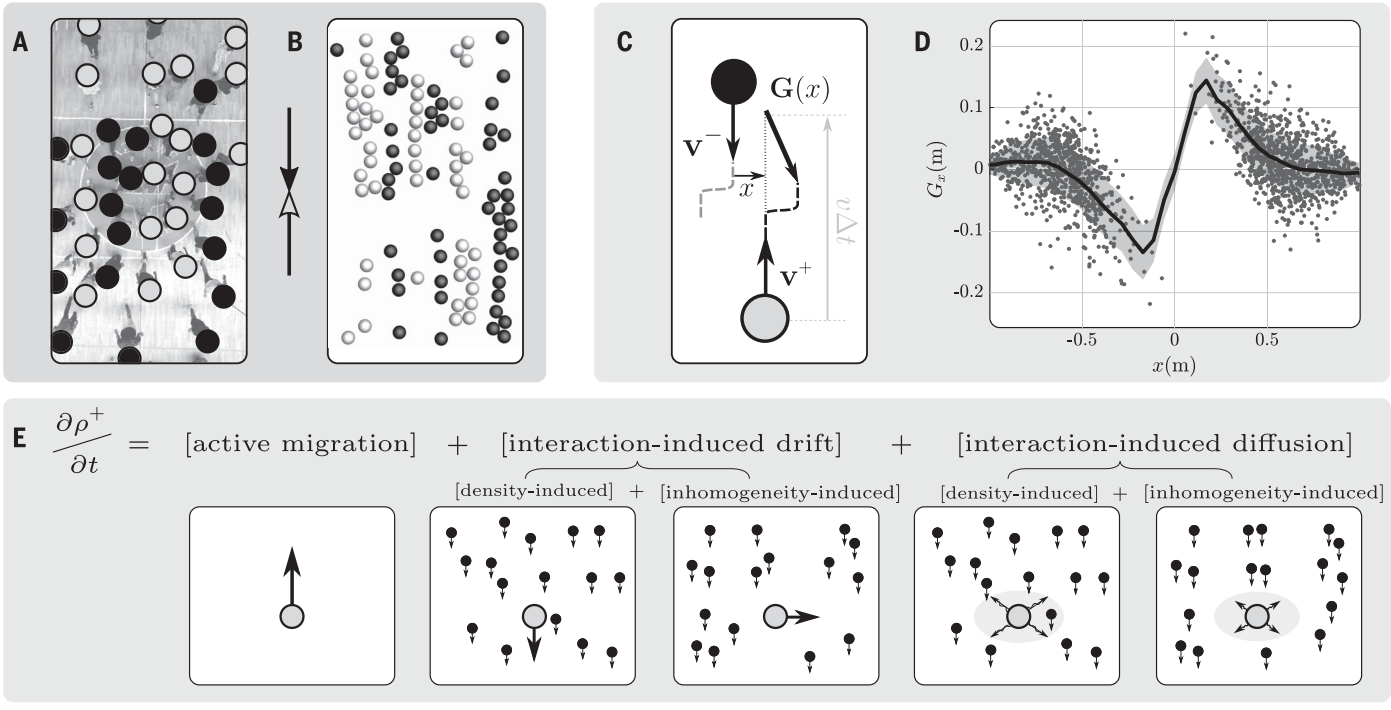


Fig. 1. Lane formation examples and kinetic model. (**A** and **B**) Lane formation examples. Bidirectional pedestrian flow realized in a controlled experiment is shown in (A). An agent-based simulation of driven hard spheres is shown in (B). (**C** to **E**) Kinetic model. In (C), a setup sketch focusing on a "+"-type agent moving with a preferred velocity $v\boldsymbol{e}_y$ (gray circle) is shown. The presence of the "−"-type agent (black circle) moving with velocity $-v\boldsymbol{e}_y$ alters its trajectory, so instead of advancing by $v\Delta t\boldsymbol{e}_y$, its displacement within a Δt time interval is given by $v\Delta t\boldsymbol{e}_y + \boldsymbol{G}(x)$, where $\boldsymbol{G}$ is the collisional operator and x is the initial lateral offset between the two agents. In (D), the x component of the collisional displacement extracted from the pedestrian experiment [see (*43*) for inference procedure] is shown. Each dark gray circle corresponds to one interaction event, the black line is a running average of these data, and the light gray shading shows the corresponding standard deviation. In (E), a diagram explaining the hydrodynamic Eq. 3 is shown. In the linearized model, the evolution of agent density can be understood as a superposition of five processes: active migration in the preferred direction, density- and inhomogeneity-induced drift (which, in a symmetric system, act in orthogonal directions), and homogeneous and inhomogeneous diffusion.

free space, whereas the sum ranges over all agents j of the opposite type who will collide with agent i in the time period of interest. The interaction events are assumed to be much shorter than the characteristic time interval Δt, which is why we refer to them as "collisions." In some systems (e.g., pedestrians), the agents do not physically collide but instead perform collision avoidance maneuvers; the distinction is not important for our theory because both effectively modify the position of interacting agents. The direction and magnitude of this perturbation are specified by the collisional operator $\boldsymbol{G}$. We choose our coordinate system so that the preferred velocities are equal in magnitude and aligned with the y axis ($\boldsymbol{v}^{\pm} = \pm v\boldsymbol{e}_y$) and so that $\boldsymbol{G}$ can be treated as a function only of the lateral displacement $x = \left(\boldsymbol{r}_i^{\pm} - \boldsymbol{r}_j^{\mp}\right) \cdot \boldsymbol{e}_x$, where $\boldsymbol{e}$ is a unit vector in standard basis (Fig. 1C).

The function $\boldsymbol{G}$ contains all necessary information about the microscopic interactions of the model and may be specified ab initio (for example, in a simulation of hard spheres) or derived from data. For deterministic dynamics, it is a univalued function, but in thermal or stochastic systems, it should be conceptualized as a random variable (*44*, *45*). Data extracted from an experiment on human crowds are given as an example of what the empirical lateral displacement $G_x(x)$ may look like in a pedestrian system (Fig. 1D).

It should be emphasized that this kinetic description is valid in the flowing regime in which pairwise interactions dominate. For higher agent densities, the dynamics will be altered by multibody interactions, eventually leading to jamming (*23*, *41*, *42*). Propagation of chaos dictates a stochastic formulation of the evolution of agent position from a disordered initial condition. We consider the position vector of each agent as a random variable with probability density function $\rho_i^{\pm}(\boldsymbol{r}, t)$ and the corresponding agent (number) densities $\rho^{\pm} = \sum_i \rho_i^{\pm}$. Applying the central limit theorem to Eq. 1, we find that $\Delta \boldsymbol{r}_i^{\pm}$ is approximately normally distributed, with mean $\pm v\Delta t(\boldsymbol{e}_y + 2\boldsymbol{A}^{\pm})$ and variance $2v\Delta t B^{\pm}$, where drift vector $\boldsymbol{A}$ and the diffusion matrix B are obtained from the agent density via

$$\boldsymbol{A}^{\pm}[\boldsymbol{r}, \rho^{\mp}] = \int \rho^{\mp}(\boldsymbol{r} + x\boldsymbol{e}_x)\mathbb{E}\boldsymbol{G}(\pm x)\mathrm{d}x,$$
$$B^{\pm}[\boldsymbol{r}, \rho^{\mp}] = \int \rho^{\mp}(\boldsymbol{r} + x\boldsymbol{e}_x)\mathbb{E}\left[\boldsymbol{G}^T(\pm x)\boldsymbol{G}(\pm x)\right]\mathrm{d}x \quad (2)$$

Put simply if the interval Δt is sufficiently long to encompass several collisions, the motion of agents or particles behaves as a set of almost uncorrelated random walks [see section I of (*43*) for the mathematical derivation]. Summing Eq. 1 and its "−"-type counterpart over i then leads to a pair of Fokker-Planck equations for the agent densities

$$\frac{\partial \rho^{\pm}}{\partial t} \pm v\frac{\partial \rho^{\pm}}{\partial y} = \mp 2v\nabla \cdot \left(\rho^{\pm}\boldsymbol{A}^{\pm}\right) + v\nabla \cdot \left(\nabla^T \rho^{\pm} B^{\pm}\right) \quad (3)$$

where $\nabla = (\partial/\partial x, \partial/\partial y)$.

We can derive several interesting conclusions from this result. Examining Eq. 3, we find on the left-hand side terms that describe simple advection in the preferred direction of motion, whereas the terms on the right-hand side describe the effects of interagent interactions: the first corresponding to interaction-induced drift (arising from the mean collisional displacement) and the second to interaction-induced diffusion (arising from the variance of collisional displacement). In the next section, we will disentangle the roles of these terms

and explore their consequences for the emergent nucleation of lanes.

Instability of homogeneous flow

The transport Eq. 3 admits a trivial homogeneous solution $\rho^{\pm}(\boldsymbol{r}, t) \equiv \rho_0$. We interrogate the stability of this solution with respect to wave-like perturbations. No generality is lost because every perturbation can be represented as a linear combination of Fourier modes, which initially evolve independently. Linearizing and taking a Fourier transform (denoted by $\tilde{}$), we obtain

$$i\omega\tilde{\rho}^{\pm} \pm ilv\tilde{\rho}^{\pm} = \mp 2iv\rho_0\tilde{\boldsymbol{A}}(0)\cdot\boldsymbol{k}\tilde{\rho}^{\pm} \mp 2vi\rho_0\tilde{\boldsymbol{A}}(\mp k)\cdot\boldsymbol{k}\tilde{\rho}^{\mp} - v\rho_0\boldsymbol{k}^T\tilde{B}(0)\boldsymbol{k}\tilde{\rho}^{\pm} - v\rho_0\boldsymbol{k}^T\tilde{B}(\mp k)\boldsymbol{k}\tilde{\rho}^{\mp} \quad (4)$$

where ω is temporal frequency and $\boldsymbol{k} = (k, l)^T$ is the wave vector, which can also be expressed by $\lambda\boldsymbol{k} = 2\pi(\cos\theta, -\sin\theta)^T$ in terms of the perturbation wavelength λ and pattern orientation θ.

The four terms on the right-hand side of Eq. 4, which are derived from the two interaction terms in Eq. 3, allow for a more nuanced analysis of the different dynamical processes. The terms involving $\tilde{\boldsymbol{A}}(0)$ and $\tilde{B}(0)$ describe the density-induced processes, and the terms involving their frequency-dependent counterparts $\tilde{\boldsymbol{A}}(k)$ and $\tilde{B}(k)$ describe control gradient-sensing processes. For example, the first term on the right describes the density-induced drag that one group imposes on the other by its sheer presence, whereas the second term describes drag that occurs in response to spatial inhomogeneities.

For chirally symmetric interactions, the density-induced drag represents the retardation associated with each collision that acts in the y direction, and the inhomogeneity-induced drag arises as a consequence of an imbalance between left and right "turns" that leads to a net drift in the x direction. In this case, as expected, one can show that the most unstable perturbations are the lane-like undulations in the x direction [$\boldsymbol{k} = (k, 0)^T$] and the growth rate $\sigma = -\Im[\omega]$ satisfies the dispersion relation

$$\sigma(k) = v\rho_0\left[2k\left|\tilde{A}_x(k)\right| + k^2\left|\tilde{B}_{xx}(k)\right| - k^2\tilde{B}_{xx}(0)\right] \quad (5)$$

Thus, we recover the two previously conjectured mechanisms of lane nucleation: gradient-induced drift proportional to $\left|\tilde{A}_x(k)\right|$ (*22*)

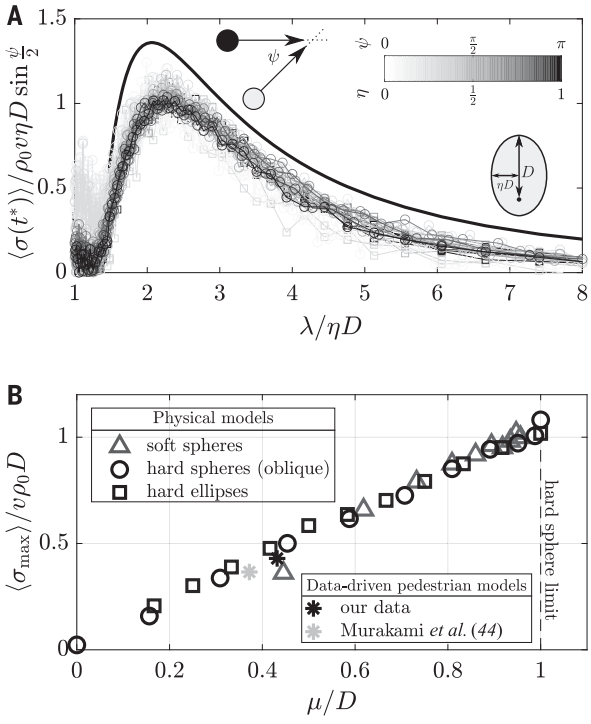


Fig. 2. Stability analysis. (**A**) The growth rate of lane-like Fourier modes at the moment of maximal linear growth t^* extracted through a Fourier transform from an ensemble of 5000 simulations of 300 hard spheres moving in a doubly periodic domain with preferred velocities subtending angle ψ (large angles correspond to black squares, as explained by the color bar at the top right), as well as hard ellipses with varying aspect ratio η (circles of varying intensity; see inset). The empirically measured growth rate shows an agreement with the theoretical dispersion relation with a global maximum for $\lambda_{max} \approx 2D$. The simulation details can be found in section IV of (*43*). (**B**) The maximal growth rate σ_{max} can also be approximated using a simple heuristic scaling $\sigma_{max} \propto v\rho_0\mu$.

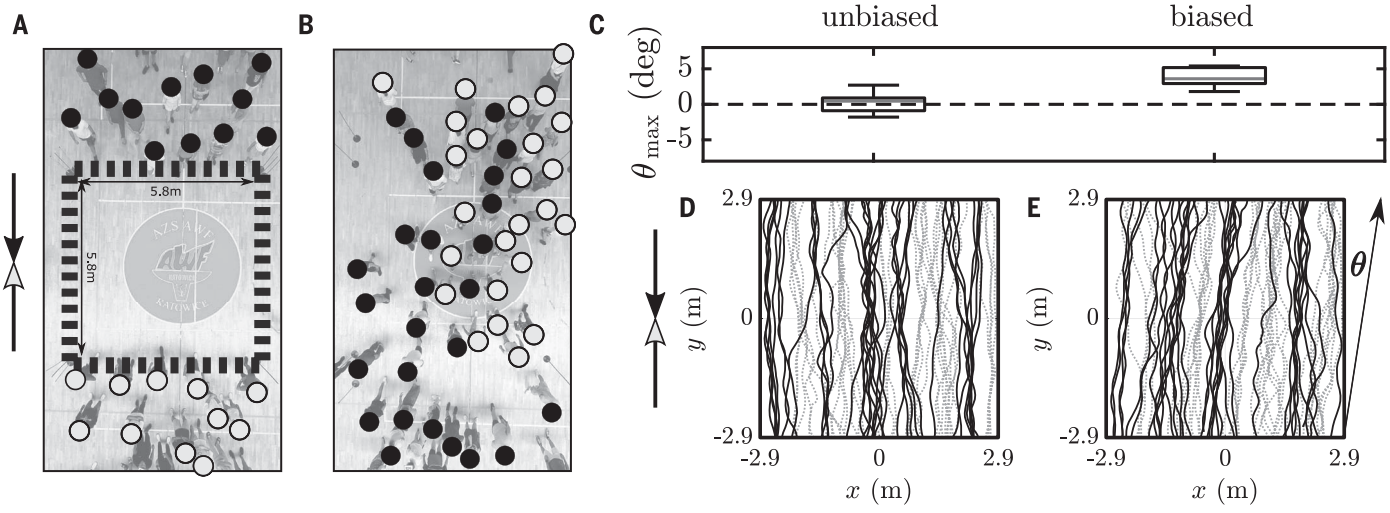


Fig. 3. Chiral intractions. (**A**) Experimental setup with participants gathering behind the starting line of the square 5.8-m–by–5.8-m arena. The experiment was repeated in two sessions, one with 60 participants and one with 73. In total, 10 benchmark trials and 7 trials with explicitly biased dodging maneuvers were recorded. (**B**) When a "pass on the right" rule is imposed, pedestrians form tilting lanes. (**C**) The tilt can be systematically detected as a dominant pattern angle through Fourier analysis of the pedestrian trajectories. The tilting trend is reproducible across the repeated trials. (**D**) Example pedestrian trajectories for unbiased interactions. (**E**) Example pedestrian trajectories from the experiment, with biased dodging maneuvers confirming the tilt.

and gradient-induced diffusion controlled by $|\tilde{B}_{xx}(k)|$ (*11*, *23*). We further discover the regularizing role of density-induced diffusion proportional to $\tilde{B}_{xx}(0)$, which prevents an "ultraviolet catastrophe" in the shortwave limit $k \to \infty$ and instead induces a cut-off wave number k_{cut} such that high-frequency oscillations are fully suppressed.

Considering the prototypical case of excluded volume interactions of hard spheres with diameter D is instructive. As calculated in (*43*), the hard-sphere dispersion relation has the analytic form

$$\sigma(k) = v\rho_0\left[-\frac{D^3}{6}k^2 + 3D - \frac{3\sin(Dk)}{k}\right] \quad (6)$$

We plot the dispersion relation as a function of wavelength λ (Fig. 2A), and it obtains a unique maximum at characteristic wavelength $\lambda_{\max} \approx 2.07D$, with a sharp cutoff to the left at $\lambda_{\mathrm{cut}} \approx 1.34D$ and an algebraic tail to the right.

Numerical simulations reveal the same characteristic shape of the dispersion relation after appropriate rescaling for hard ellipses and hard spheres with nonparallel preferred velocity $\boldsymbol{v}^{\pm}$ (Fig. 2A). The growth of lanes exhibits some strongly nonlinear characteristics, in that the logarithmic derivative of a given average Fourier mode $\langle\sigma(k)\rangle = \langle\tilde{\dot{\rho}}^{\pm}\rangle/\langle\tilde{\rho}^{\pm}\rangle$ is not constant in time [section IV of (*43*)]. Nevertheless, at the "moment of maximal linear growth" t^* that dominates lane formation, such that $\langle\sigma_{\max}\rangle = \langle\sigma\rangle[(k^*, 0), t^*] = \max_{k,t}\langle\sigma\rangle[(k, 0), t]$, we find close agreement between $\langle\sigma(k)\rangle$ and the theoretically predicted growth rate (Eq. 5). We confirm this relationship for excluded volume models of ellipses and spheres (Fig. 2A), but a similar agreement has been found for other deterministic and stochastic agent-based models [see sections IV and V of (*43*)]. These models include a data-driven event-based simulation that bootstraps the experimental displacement data that is displayed (Fig. 1D) to generate a large number of "in silico" pedestrian experiments with the same displacement statistics.

Dimensional analysis of Eq. 5 shows that the k-dependent terms act to select a characteristic length scale of lane nucleation, itself emerging from the action of the displacement operator. This length scale can be closely approximated by the square root of the L^1 norm of the x component, $\mu = \sqrt{2\|G_x\|_1}$. We show that the rate of lane nucleation measured from simulations is robustly predicted by the simple expression $\sigma_{\max} \approx v\rho_0\mu$ (Fig. 2B). The first two terms of this expression determine the rate of collisions, whereas μ controls the size the lateral displacement per collision, which might vary as a result of the softness of particles or the angle of approach.

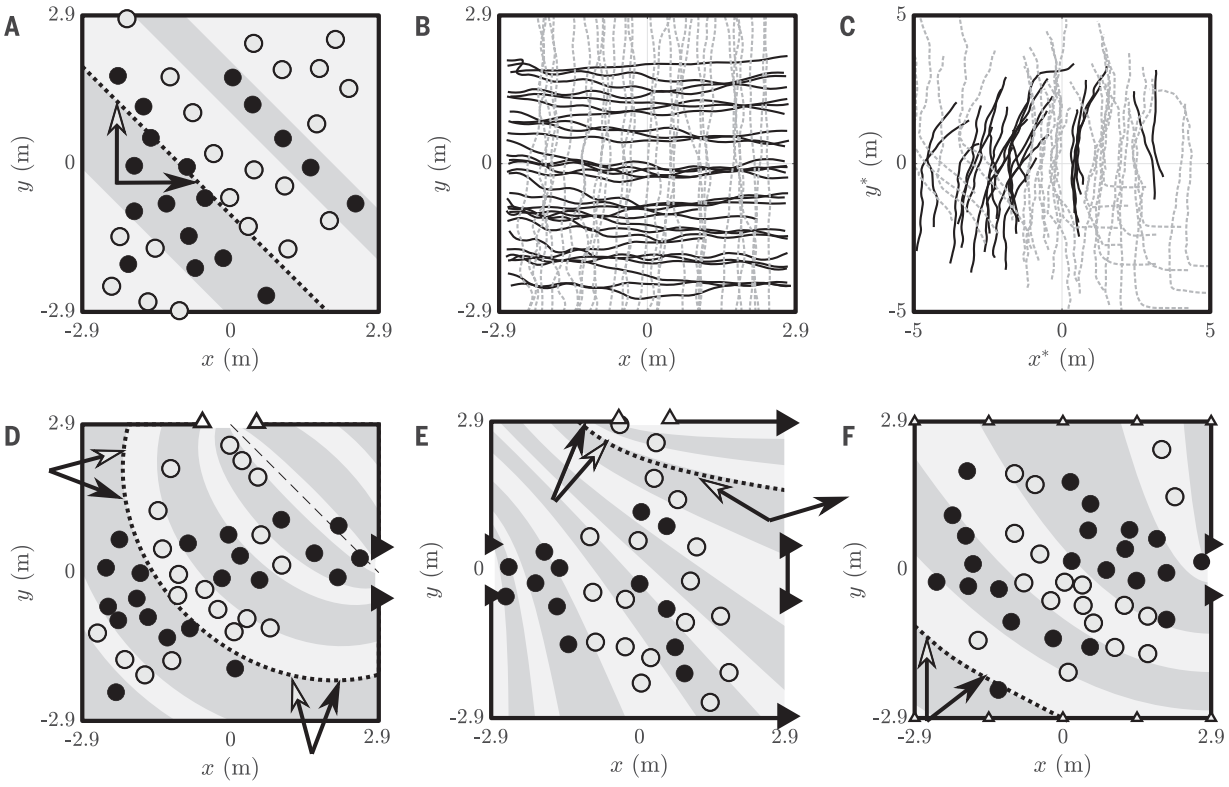


Fig. 4. Complex flows. (**A**) Two groups of agents with perpendicular target velocities form diagonal lanes that migrate with the average velocity. The arrows highlight that the lanes are expected to form along the direction of differential velocity, indicated by the dotted line, as well as the dark gray–and light gray–shaded regions. The circles represent the positions of pedestrians that are crossing the experimental arena (Fig. 3A) as two perpendicular streams. (**B**) Pedestrian trajectories in the lab frame cross at 90°. (**C**) The same trajectories, as viewed in an appropriately rotated reference frame moving with a speed of 0.7 m/s (approximating the average velocity of the two groups), reveal the laning pattern. (**D**) For a more complex scenario, in which the two groups head toward two point targets, the laning instability is triggered along ellipses. This can be proved by using the fact that lanes emerge along the direction of differential velocity (see arrows) as well as the reflective properties of an ellipse. As before, the black and light gray circles indicate the pedestrian positions observed in an experiment, which show segregation consistent with the theoretically predicted elliptical lines, indicated by the shaded regions. The black and light gray triangles indicate the gates targeted by the respective groups. (**E**) When one of the groups (black circles) moves away from a particular point, the laning lines form a family of hyperbolas. In the experiment, the "repulsive point" is replaced with a narrow entry gate (indicated with black triangles on the left). (**F**) When one group (light gray circles) has a preferred direction and the other (black circles) has a point target, the incipient lanes form along parabolas.

Chiral interactions

We emphasize that the fact that the most unstable perturbations of the homogeneous state are aligned with the differential velocity is only the case for chirally symmetric interactions. If the collisional displacement $\boldsymbol{G}(x)$ is not symmetric about zero, one can show that the maximal growth rate is attained at

$$\theta_{\max} \approx \tan^{-1}\left[2\rho_0 \tilde{A}_x(0)\right] \tag{7}$$

The inclination of incipient lanes is driven through density-induced drift, as evident from this expression, which itself can provide an alternative instability mechanism (*30*). Indeed, if the interactions are chirally biased, the very presence of the other agents will induce drift and agent segregation. This should be contrasted with the symmetric case, where the drift is gradient-induced, that is, it relies on the imbalance of left and right turns that are associated with encountering more agents of the opposite type on the left or right.

The qualitative prediction of tilting lanes has been confirmed in a controlled experiment with pedestrians moving across a square arena (Fig. 3). When the participants were divided into two groups and asked to cross to the other side, they formed lanes parallel to the direction of motion (movie S1). However, when the experiment marshals instructed the participants to repeat the same exercise while obeying a "pass on the right" traffic rule, the crowds spontaneously segregated into lanes with marked inclination (movie S2). A nontrivial tilting angle has also been confirmed in a quantitative manner with a Fourier analysis of the trajectory data from repeated trials. The magnitude of the tilt is consistent with the properties of the empirical collisional operator [section VII of (*43*)]. The validity of Eq. 7 has also been confirmed by agent-based simulations of "frictional rotating disks" [see section IV of (*43*)]. It is also worth noting that if the lane orientation and the direction of motion are misaligned, collisions do not cease once the lanes are formed and the agents are forced to increase the lengths of their paths. None of these effects are desired in pedestrian traffic, which suggests that, in crowd management, the "pass on the right" rule should be used with caution.

Complex flows

Our theoretical framework is also easily extended to agents with generalized preferred velocity. Indeed, this case can be reduced to the system with $\boldsymbol{v}^{\pm} = \pm v\boldsymbol{e}_y$ by performing a local orthogonal coordinate transformation into a frame moving with the average speed $\frac{1}{2}[\boldsymbol{v}^{+} + \boldsymbol{v}^{-}]$ and the y axis aligned with the differential velocity $\boldsymbol{v}^{+} - \boldsymbol{v}^{-}$. Consequently, in this case, lanes emerge along the direction of the differential velocity, grow at a theoretical rate proportional to $v = \frac{1}{2}|\boldsymbol{v}^{+} - \boldsymbol{v}^{-}|$ (compare with Eq. 5), and shift with the average velocity (*46*). We show that this theoretical estimate of lane growth rate is consistent with hard-sphere simulations because we vary the angle ψ between the two preferred directions (Fig. 2A).

The formation of traveling laning patterns has also been observed in the experiment with two groups of pedestrians moving in perpendicular directions; in this case, the lanes are inclined at 45° (Fig. 4A and movie S3) (*47*). We confirm that even though the diagonal organization appears more complex, it corresponds to a "simple" stationary laning pattern in the appropriate frame of reference (Fig. 4C).

The observation that in symmetric systems lanes form along the direction of differential velocity can be used to analyze lane nucleation in more general cases where the preferred velocity varies in space. This may occur, for example, when the agents proceed with equal speeds toward two distinct point targets $\boldsymbol{f}^{\pm}$, where the lane formation is triggered along ellipses with foci in $\boldsymbol{f}^{\pm}$ (Fig. 4D). When $\boldsymbol{f}^{+}$ is a target but $\boldsymbol{f}^{-}$ is a point of repulsion, lanes nucleate along hyperbolas with foci $\boldsymbol{f}^{\pm}$ (Fig. 4E). Finally, when the "+" group has a target direction $\boldsymbol{v}^{+}$ but the "–" group has a target point $\boldsymbol{f}^{-}$, the pattern curves are parabolas with a focus in $\boldsymbol{f}^{-}$ and a directrix perpendicular to $\boldsymbol{v}^{+}$ (Fig. 4F). All of these facts can be easily proved using the reflective properties of conic sections. In more sophisticated scenarios, for example, when the two speeds are not equal, computing the laning curves may require numerical integration. Importantly, these systems do not possess a stable steady state, so the local onset of lane nucleation does not progress to nonlinear growth and saturation. In particular, the accumulation of agents near the attractive points will lead to an increase in density and eventually to the breakdown of our theory as multibody interactions dominate.

Despite these theoretical limitations, we have observed curved lane nucleation in a suite of controlled pedestrian experiments that mimic the three canonical scenarios (Fig. 4, D to F, and movies S4 to S6). In the experiment, the point targets are associated with narrow exit gates and the point repeller is approximately realized with a narrow entry gate. To avoid excessive crowding, the entrance and exit gates were appropriately wide (1 m) relative to the throughput of pedestrians; although higher pedestrian densities were observed near gates, encounters between agents of opposite types were distributed across the center of the experimental domain [(*43*); fig. S25]. In each case, the directions of the nucleating lanes were found to be consistent with the theoretical prediction. We also show that conic section patterns are observed in numerical simulations [section VI of (*43*)], where the average laning direction over a large number of replicates can be extracted in a quantitative manner.

Discussion

We considered the complex flow of two groups of agents driven toward different targets. Most importantly, we have introduced a simple kinetic framework that directly relates the individual behavior agents to their collective dynamics, thereby bridging the gap between the exact agent-based studies (*38*, *48*, *49*) and heuristic continuum models (*2*, *19*, *29*, *30*, *50*). Our analysis systematizes different previously proposed mechanisms of lane formation in dilute active mixtures (*11*, *22*, *23*, *30*) and provides quantitative estimates for the propensity of lane nucleation that is induced by binary collision-like interactions. Our theory also generates predictions about the nontrivial orientation of lanes in chirally biased systems, as well as curved lanes in complex flows with spatially varying directions of motion. All these have been confirmed by numerical simulations and in a controlled pedestrian experiment with human participants. The latter validation step is particularly important because it shows that our model correctly describes the characteristics of lane nucleation in real-world systems with complex interaction laws.

REFERENCES AND NOTES

1. M. Isobe, T. Adachi, T. Nagatani, *Physica A* **336**, 638–650 (2004).
2. D. Helbing, L. Buzna, A. Johansson, T. Werner, *Transport. Sci.* **39**, 1–24 (2005).
3. J. Zhang, W. Klingsch, A. Schadschneider, A. Seyfried, *J. Stat. Mech.* **2012**, P02002 (2012).
4. J. Bamberger *et al.*, in *Traffic and Granular Flow '13*, M. Chraibi, M. Boltes, A. Schadschneider, A. Seyfried, Eds. (Springer, 2015), pp. 137–144.
5. S. Cao, A. Seyfried, J. Zhang, S. Holl, W. Song, *J. Stat. Mech.* **2017**, 033404 (2017).
6. C. Feliciani, K. Nishinari, *Phys. Rev. E* **94**, 032304 (2016).
7. N. Guo, Q. Y. Hao, R. Jiang, M. B. Hu, B. Jia, *Transp. Res. Part C Emerg. Technol.* **71**, 63–85 (2016).
8. N. Guo, H. X. Liu, R. Jiang, B. Jia, M. B. Hu, *Int. J. Mod. Phys. C* **29**, 1850069 (2018).
9. M. Moussaïd *et al.*, *PLOS Comput. Biol.* **8**, e1002442 (2012).
10. M. E. Leunissen *et al.*, *Nature* **437**, 235–240 (2005).
11. T. Vissers *et al.*, *Soft Matter* **7**, 2352–2356 (2011).
12. K. R. Sütterlin *et al.*, *Phys. Rev. Lett.* **102**, 085003 (2009).
13. C. R. Du *et al.*, *New J. Phys.* **14**, 073058 (2012).
14. R. Jose, L. Santen, *Phys. Rev. Lett.* **124**, 198103 (2020).
15. S. Nowak, A. Schadschneider, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **85**, 066128 (2012).
16. Y. Sun, *Physica A* **549**, 124295 (2020).
17. M. Burger, S. Hittmeir, H. Ranetbauer, M. T. Wolfram, *SIAM J. Math. Anal.* **48**, 981–1005 (2016).
18. H. Ohta, *Europhys. Lett.* **99**, 40006 (2012).
19. J. Chakrabarti, J. Dzubiella, H. Löwen, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **70**, 012401 (2004).
20. J. Cristín, V. Méndez, D. Campos, *Sci. Rep.* **9**, 18488 (2019).
21. J. Delhommelle, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **71**, 016705 (2005).
22. D. Helbing, T. Vicsek, *New J. Phys.* **1**, 13 (1999).
23. K. Klymko, P. L. Geissler, S. Whitelam, *Phys. Rev. E* **94**, 022608 (2016).
24. A. Poncet, O. Bénichou, V. Démery, G. Oshanin, *Phys. Rev. Lett.* **118**, 118002 (2017).
25. M. Rex, H. Löwen, *Eur. Phys. J. E* **26**, 143–150 (2008).
26. T. Xiong, P. Zhang, S. C. Wong, C. W. Shu, M. P. Zhang, *Chin. Phys. Lett.* **28**, 108901 (2011).
27. H. Kuang, X. Li, T. Song, S. Dai, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **78**, 066117 (2008).
28. C. Totzeck, *Kinetic Relat. Models* **13**, 1219–1242 (2020).
29. J. Chakrabarti, J. Dzubiella, H. Löwen, *Europhys. Lett.* **61**, 415–421 (2003).
30. D. Helbing, *Complex Syst.* **6**, 391–415 (1992).
31. A. Bricard, J.-B. Caussin, N. Desreumaux, O. Dauchot, D. Bartolo, *Nature* **503**, 95–98 (2013).
32. N. Bain, D. Bartolo, *Science* **363**, 46–49 (2019).

33. N. Bain, D. Bartolo, *Nat. Commun.* **8**, 15969 (2017).
34. A. Einstein, *Investigations on the Theory of the Brownian Movement* (Dover Publications, 1956).
35. X. Jia, C. Feliciani, D. Yanagisawa, K. Nishinari, *Physica A* **531**, 121735 (2019).
36. M. Moussaïd *et al.*, *Proc. Biol. Sci.* **276**, 2755–2762 (2009).
37. C. Scholz, M. Engel, T. Pöschel, *Nat. Commun.* **9**, 931 (2018).
38. I. Karamouzas, B. Skinner, S. J. Guy, *Phys. Rev. Lett.* **113**, 238701 (2014).
39. Q. Xu, M. Chraibi, A. Seyfried, *Transp. Res. Part C Emerg. Technol.* **133**, 103464 (2021).
40. J. Wang *et al.*, *Physica A* **500**, 106–120 (2018).
41. M. Muramatsu, T. Irie, T. Nagatani, *Physica A* **267**, 487–498 (1999).
42. C. Reichhardt, C. J. O. Reichhardt, *Soft Matter* **14**, 490–498 (2018).
43. See materials and methods.
44. H. Murakami, C. Feliciani, Y. Nishiyama, K. Nishinari, *Sci. Adv.* **7**, eabe7758 (2021).
45. A. Corbetta, J. A. Meeusen, C. M. Lee, R. Benzi, F. Toschi, *Phys. Rev. E* **98**, 062310 (2018).
46. J. Dzubiella, H. Löwen, *J. Phys. Condens. Matter* **14**, 9383–9395 (2002).
47. P. Mullick *et al.*, *PLOS Comput. Biol.* **18**, e1010210 (2022).
48. D. Helbing, P. Molnár, *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **51**, 4282–4286 (1995).
49. F. Dietrich, G. Köster, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **89**, 062801 (2014).
50. J. Toner, Y. Tu, *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **58**, 4828–4858 (1998).
51. K. A. Bacik, B. S. Bacik, T. Rogers, Dataset for "Lane nucleation in complex active flows." University of Bath Research Data Archive (2023); https://doi.org/10.15125/BATH-01242.

ACKNOWLEDGMENTS

We thank colleagues from the Department of Human Motor Behavior at the Academy of Physical Education in Katowice—in particular, G. Sobota, W. Marszałek, and M. Pawłowski—for their help in conducting the experiment. K.A.B. and T.R. thank E. Gavagnin and J. Meyer for their advice on experiments involving AprilTags and humans, respectively, as well as B. Hayward for useful discussions. We also thank A. Souslov for proofreading the manuscript. We acknowledge the University of Bath's Research Computing Group (https://doi.org/10.15125/b6cd-s854) for their support in this work. **Funding:** K.A.B. and T.R. acknowledge support from Engineering and Physical Sciences Research Council grant number EP/V048228/1. The cost of the experiment was covered by the Academy of Physical Education in Katowice. **Author contributions:** K.A.B. performed the theoretical analysis and numerical simulations, designed the experiments, and drafted the manuscript. B.S.B. designed and conducted the experiments. T.R. conceived the study, performed the theoretical analysis, and drafted the manuscript. All authors approved the final text of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data from the pedestrian experiment is available at the University of Bath Research Data Archive (*51*). The research protocol of the pedestrian experiment was approved by the Bioethics Committee for Scientific Research at the Academy of Physical Education in Katowice (no. 1/2022/6/23) and met the ethical standards of the Declaration of Helsinki.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add8091
Materials and Methods
References (*52*, *53*)
Movies S1 to S6

Submitted 11 July 2022; accepted 2 February 2023
10.1126/science.add8091

VOCALIZATION

Toothed whales use distinct vocal registers for echolocation and communication

Peter T. Madsen[1], Ursula Siebert[2], Coen P. H. Elemans[3]*

Echolocating toothed whales (odontocetes) capture fast-moving prey in dark marine environments, which critically depends on their ability to generate powerful, ultrasonic clicks. How their supposedly air-driven sound source can produce biosonar clicks at depths of >1000 meters, while also producing rich vocal repertoires to mediate complex social communication, remains unknown. We show that odontocetes possess a sound production system based on air driven through nasal passages that is functionally analogous to laryngeal and syringeal sound production. Tissue vibration in different registers produces distinct echolocation and communication signals across all major odontocete clades, and thus provides a physiological basis for classifying their vocal repertoires. The vocal fry register is used by species from porpoises to sperm whales for generating powerful, highly air-efficient echolocation clicks.

Toothed whales, the odontocetes, have access to rich marine food resources down to depths of 2000 m and achieve a biomass turnover larger than that of human fisheries combined (*1*). The key to this success is their ability to locate, track, and capture fast-moving prey underwater in complete darkness at depths of, routinely, >100 m using echolocation, a feat that critically depends on the production of short, powerful, ultrasonic echolocation clicks (hereafter clicks) at rates >300 clicks per second (*2*). Paradoxically, odontocetes are thought to produce clicks with an airflow-driven sound source in their nose (*3*), but how they can produce clicks with less than 10% remaining air volume and pressure-collapsed lungs at depths beyond 100 m is not understood. The same source is thought to be sufficiently plastic to also produce rich, learned acoustic repertoires essential for mediating complex social interactions (*4*). However, if so, how the same source can generate such a diversity of acoustic signals to serve both echolocation and communication remains unknown.

Echolocation clicks are made nasally, not laryngeally

The location of the odontocete sound source was initially hypothesized to be the larynx, but was later designated to the so-called phonic lips in the nose (*5*) on the basis of three lines of evidence: (i) laryngeal muscles are inactive during clicking (*6*), (ii) air is pressurized and recycled in between air sacs surrounding the nasal passages (*7*, *8*), and (iii) acoustic triangulation localizes the click source 20 to 70 mm below the blowhole (*9*). Determining the odontocete sound source location requires quantification of phonic lip motion dynamics with high spatiotemporal resolution to link tissue motion dynamics to sound generation. However, because of experimental limitations, imaging sound-producing events in the odontocete nose at sufficient speed in vivo has remained very challenging (*3*).

To test the hypothesis that phonic lips are the odontocete click source, we imaged their motion at 7200 frames per second (fps) using an endoscope while simultaneously measuring air pressure below (p_{sub}) and above (p_{supra}) the phonic lips during click production in trained Atlantic bottlenose dolphins (*Tursiops truncatus*, $n = 2$) and Harbor porpoises (*Phocoena phocoena*, $n = 3$), in vivo (Fig. 1B) (*10*). Dolphins and porpoises produced clicks exclusively during prolonged bouts of increased p_{sub} (Fig. 1D), and 98% were produced above $p_{sub} = 3.63 \pm 0.14$ kPa and 2.12 ± 1.56 kPa, respectively (Fig. 1E), which were consistent with earlier measurements (*3*, *6*, *11*).

Phonic lips and more deeply situated bilateral nasal plugs (Fig. 1B) were clearly visible during inspiration, but during pressurization, dorsal airsac walls moved rostrally and occluded direct views (Fig. 1F). We observed click-associated tissue motion with a delay of 1.0 ± 1.3 ms in dolphins ($n = 2$) and 1.5 ± 0.36 ms in porpoises ($n = 3$) after click emission at the melon (Fig. 1, G to I and movies S1 and S2). Therefore, we likely did not observe phonic lips directly, but mucosal waves traveling rostrally over the nasal surface. Mucosal waves are common during tissue vibration–induced sound production (*12*, *13*). Combining motion latency with a mucosal wave speed of ~1 m/s (*12*) predicts the motion source to be ~1 to 2 mm below the observed slit, which is consistent with the anatomical location of phonic lips. Thus, our data directly links nasal tissue motion to sound generation and confirms the nasal biosonar sound source.

Echolocation clicks are produced by colliding phonic lips during flow-induced voiced sound production

Identifying the physiological mechanism that drives phonic lip motion is crucial for understanding the parameters that determine acoustic output, from immediate motor control to morphological changes over evolutionary time scales. Two mechanisms have been suggested for toothed whale sound production: superfast muscle actuation (*14*, *15*) and air flow–induced self-sustained vibration (*3*, *6*, *7*, *11*). The latter is consistent with the myoelastic-aerodynamic theory of sound production for laryngeal and syringeal sound production (*12*, *13*). Myoelastic-aerodynamic sources require air pressurization and volume flow, which are complicated by the highly reduced air volumes with diving depth and should thus severely constrain the functional depth range of toothed whale sound production. Alternatively, phonic lip motion actuated by superfast muscles would not be depth dependent and would thus circumvent this constraint. Indeed, odontocete genomes imply the plausibility of superfast muscles to power rapid phonic lip motion (*14*) and motor activity may precede individual clicks (*15*). Although weak superfast muscles may not be capable of powering heavy phonic lip motion (*16*), they may trigger catch-release or stridulation mechanisms as in fish and aquatic frogs (*17*).

To test whether a myoelastic-aerodynamic source or muscle contraction powers click production, we imaged phonic lips during sound production in an in vitro preparation of the nasal complex in dead harbor porpoises, thereby excluding neural control or muscle action (Fig. 2) (*10*). At nasal threshold pressures >5 kPa, we induced emission of click sequences in six specimens by phonic lip adduction. Imaging of phonic lip kinematics (4000 fps, 8 bits) showed that anterior and posterior phonic lips undergo flow-induced, self-sustained oscillations during click emission (Fig. 2, B to D). Click production was tightly associated with phonic lip collision (delay of 74 ± 204 μs), not opening (delay of 1.70 ± 1.16 ms) (Fig. 2, E and F). Higher-precision imaging (10,000 fps, 12 bits) showed that phonic lip collisions occurred only 57 ± 103 μs ($n = 2370$) prior to click production (Fig. 2, E and F, and movie S3). Thus, phonic lip collision, not opening, causes tissue excitation that generates pressure waves emanating as clicks from the melon, which supports previous speculations (*3*, *15*). A click-triggered, averaged image upsampled to 100,000 fps suggests that the phonic lip leading edge collides before the trailing edge (Fig. 2G). This is consistent with a caudocranial travelling

[1]Zoophysiology, Department of Biology, Aarhus University, 8000 Aarhus, Denmark. [2]Institute for Terrestrial and Aquatic Wildlife Research, University of Veterinary Medicine Hannover, 25761 Büsum, Germany. [3]Sound Communication and Behavior Group, Department of Biology, University of Southern Denmark, 5230 Odense M, Denmark.
*Corresponding author. Email: coen@biology.sdu.dk

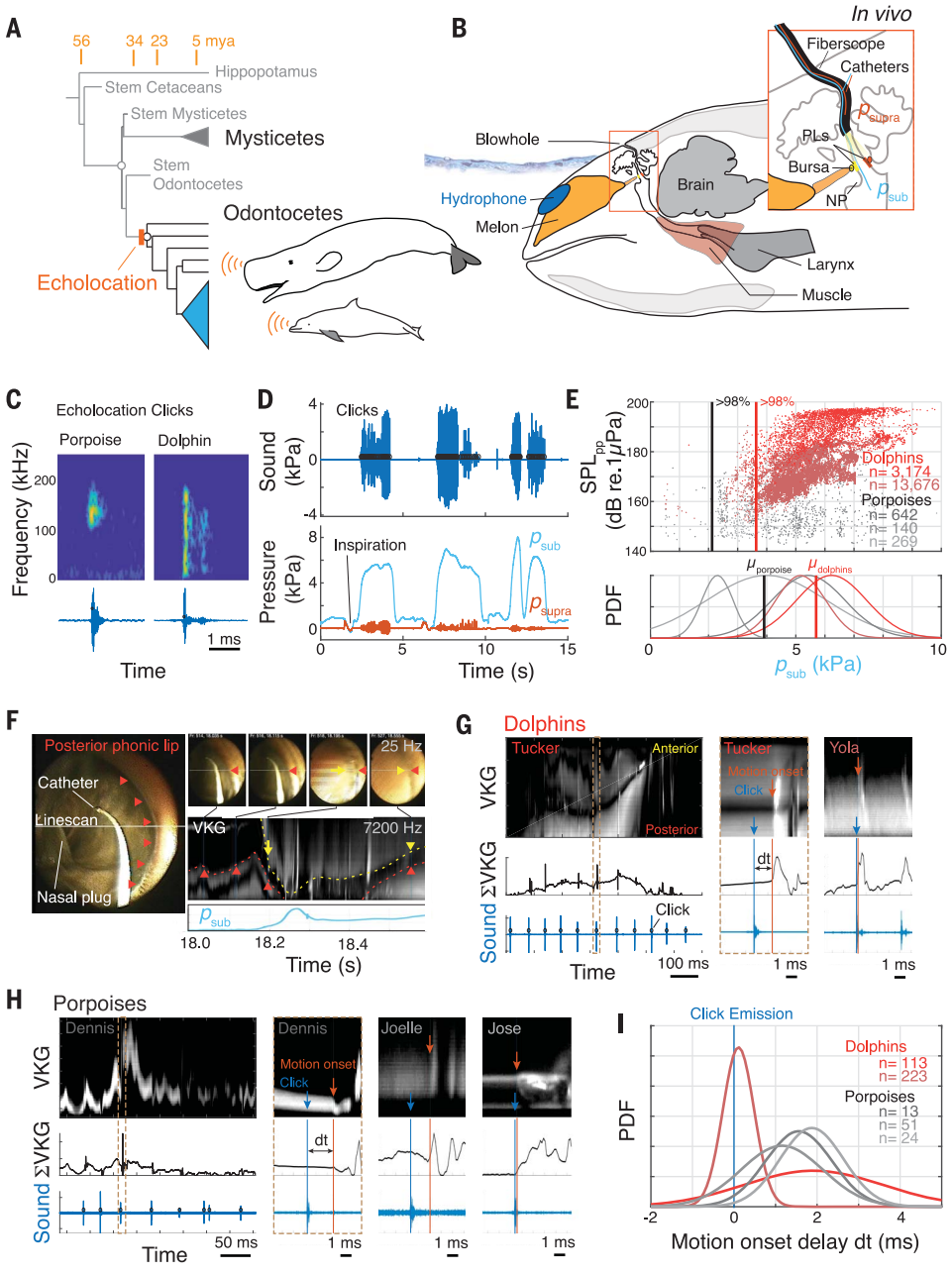


Fig. 1. Odontocete echolocation clicks are generated in nasal passages. (**A**) Odontocetes evolved the ability to echolocate (*32*). (**B**) In vivo recording schematic with hydrophone, endoscope, and pressure catheters placed above (p_{supra}) and below (p_{sub}) the phonic lips (PLs). NP, nasal plug (*23*). (**C**) Harbor porpoise and bottlenose dolphin echolocation clicks are radiated during (**D**), p_{sub} increase. (**E**) Probability density functions (PDF) per individual. The mean (μ) p_{sub} values per species during click emission were 5.70 ± 0.75 kPa and 3.91 ± 1.57 kPa for dolphin and porpoises, respectively. (**F**) Full-frame color video (25 fps) during inspiration (left), showing the posterior PLs (red arrowheads), p_{sub} pressure catheter, and location of the 7200 fps videokymography (VKG) linescan. During p_{sub} increase (right), anterior PLs (yellow arrowheads) quickly move posterior. (**G**) Example VKG, summed VKG intensity, and hydrophone signal in the nares of two dolphins and (**H**) three porpoises with click (blue) and motion onset (orange vertical lines). (**I**) Distributions of click-associated motion delay (dt) observations per individual.

mucosal wave, critical to sustaining vibrations in myoelastic-aerodynamic sound sources (*12*, *13*). Because natural porpoise clicks can be produced in vitro without motor control, clicks are produced by phonic lip oscillations conforming to the myoelastic-aerodynamic theory and not by superfast muscle contraction–induced motion. Thus, odontocetes do not control the timing and level of individual clicks, but modulate click rates and levels by motor control of phonic lip tension and nasal pressure.

The odontocete acoustic repertoires arise from distinct vocal registers

Next to clicks, many odontocetes produce rich repertoires of lower-intensity, lower-frequency communication sounds described qualitatively as bursts, grunts, and whistles (*18*). These sounds are presumably also produced by phonic lips (*6*, *9*), but it remains puzzling how the same structures generate such diversity of acoustic signals. A characteristic of laryngeal myoelastic-aerodynamic sound production is that vocal folds exhibit self-sustained oscillation in distinct patterns, so-called registers, or laryngeal mechanisms (*19*) (Fig. 3). However, it is unknown whether odontocete phonic lips can vibrate in different registers and, if so, if these registers could generate the diversity of observed acoustic signals.

In humans, at least three registers are recognized (Fig. 3A) (*19*, *20*). In vibrational register M0 (vocal fry), both vocal fold cover and body are slack, leading to the lowest frequencies with a short open phase of the vibratory cycle [open quotient (OQ) of 0 to 0.4] (*19*). Glottal flow is low and vocal fold acceleration and tracheal sound are pulsatile (*19*, *20*). In register M1(chest), vocal folds are lengthened and the vocal fold body is stiffer than the cover, leading to higher vibration frequencies at an OQ between 0.3 and 0.8. Glottal flow is low and vocal fold acceleration and tracheal sound waveforms are triangular and sinusoidal, respectively (*19*, *20*). In vibrational register M2 (falsetto), vocal folds are lengthened further with both body and cover stiff, leading to higher frequencies

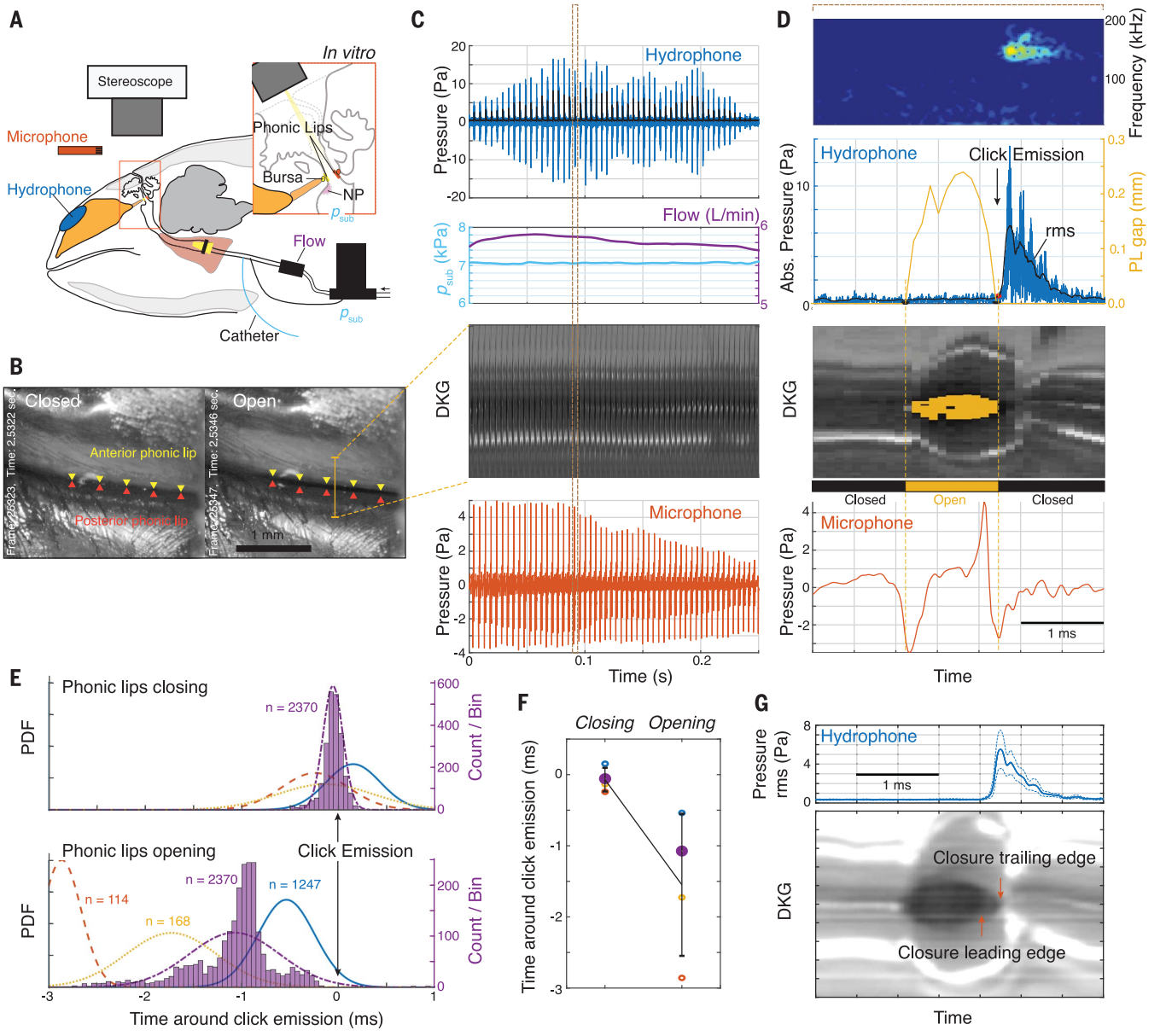


Fig. 2. Harbor porpoise echolocation clicks are produced by colliding phonic lips during flow-induced voiced-sound production. (**A**) In vitro setup to induce sound production with removed nasal plug (NP). (**B**) Still image of open and closed right PLs filmed at 10,000 fps. Vertical yellow line indicates Digital Kymogram (DKG, right) location. (**C**) Signal overview and detail (dashed rectangle) (**D**) during induced click. (**E**) Distributions of PL closing (top) and opening (bottom) events relative to click emission from the melon show that (**F**) closing is closely associated with phonic lip collision, not opening. Colors indicate individuals, with specimen P24152 (10,000 fps, 12 bits) in purple. (**G**) Averaged DKG (bottom) of 30 upsampled (100,000 fps) DKG segments aligned to click emission (top).

and an OQ between 0.5 and 0.95 (*19*). Glottal flow is high and vocal fold acceleration and tracheal sound waveforms are nearly sinusoidal with strong fundamental frequency (f_o) (*19*).

Thus, at least three independent features inform a test of our hypothesis that odontocetes vocalize in different registers. (i) Anatomically, layered phonic lips would facilitate registers by allowing differential tension between cover and body (*19*, *21*, *22*). (ii) Acoustically, different register vocalizations should have distinct sound pressure waveforms and overlapping, increasing f_o ranges (Fig. 3D) (*20*). (iii) Physiologically, different register vibrations should have increasing OQ values (*19*).

To test whether odontocete phonic lip anatomy supports register vibrations, we quantified phonic lip geometry using contrast-enhanced DiceCT (*10*). Corroborating earlier work (*5*, *23*), we observed a superficial cover layer in porpoise phonic lips with bilaterally paired deeper fat bodies called bursae (Figs. 3B), which are present in all odontocetes (*5*). We propose that bursae are functionally analogous to the vocal fold body layer and allow deep tissue rotation during paired phonic lip oscillation (Fig. 3B). Furthermore, we propose the cover layer has two morphological adaptations to optimize M0 vibration. First, increased height compared with that of laryngeal vocal folds (Fig. 3B), which reduces OQ (*22*). Second, phonic lips are covered by ridges (Fig. 3B and fig. S1) that lead to strong stiffness anisotropy in the superficial layer, but not in the body. This anisotropy improves glottal closure and the ability to maintain adductory position (*22*) against driving air pressures as high as 40 to 81 kPa (*11*).

To test whether the acoustic diversity of odontocete phonations supports different registers, we compiled acoustic repertoires for representative species from all major clades (*10*) (table S1). The repertoires indeed comprise distinct phonation types of which the waveforms are pulsatile ("clicks"), harmonic ("bursts"), or sinusoidal ("whistles"), consistent with the vocal fold acceleration waveform of registers M0 to M2 (Fig. 3C). Furthermore, the vocalization types have overlapping, increasing f_o ranges (Fig. 3, D and E, and table S1).

Vocal fry powers echolocation

Next, we tested whether phonic lip vibration during clicks is consistent with the M0 register and measured clicks OQ values in free-swimming odontocetes (*10*) (Fig. 4, A to C). Because we could not film phonic lip kinematics in the open ocean, we leveraged the causal biomechanical interpretation of our in vitro and in vivo data (Figs. 2 and 3). We measured the timing of acoustic signatures of phonic lip opening and closing events on microphone and hydrophone signals during click emission in trained porpoises and dolphins (*10*) (Fig. 4, A to C), and on hydrophone signals of acoustic biologging tags during dives in free-swimming Bottlenosed dolphins, false killer whales, and sperm whales, down to depths of 1800 m (Fig. 4B and fig. S2). For clicks, the OQ remained <40% in all species and increased significantly with click rate (Fig. 4E and table S2). Additionally, opening duration changed with click rates so that OQ stayed <10% (Fig. 4F). We propose that odontocetes reduce OQ to increase phonic lip acceleration, thereby increasing click source level. Thus, clicks in free swimming odontocetes are produced by phonic lip vibration in the M0 register, and their OQ is extremely low.

Lastly, we measured an OQ of 0.35 to 0.8 during burst vocalizations in a bottlenose dolphin (Fig. 4, figs. S3 and S4, and movie S4), which was consistent with M1 register vibration and distinct from clicks (M0) in the same f_o range. The similar OQ ranges for M0 and M1 vibration in human laryngeal voice production (*19*) suggest either functional convergence between independently evolved systems or a basal transition property of myoelastic-aerodynamic sources.

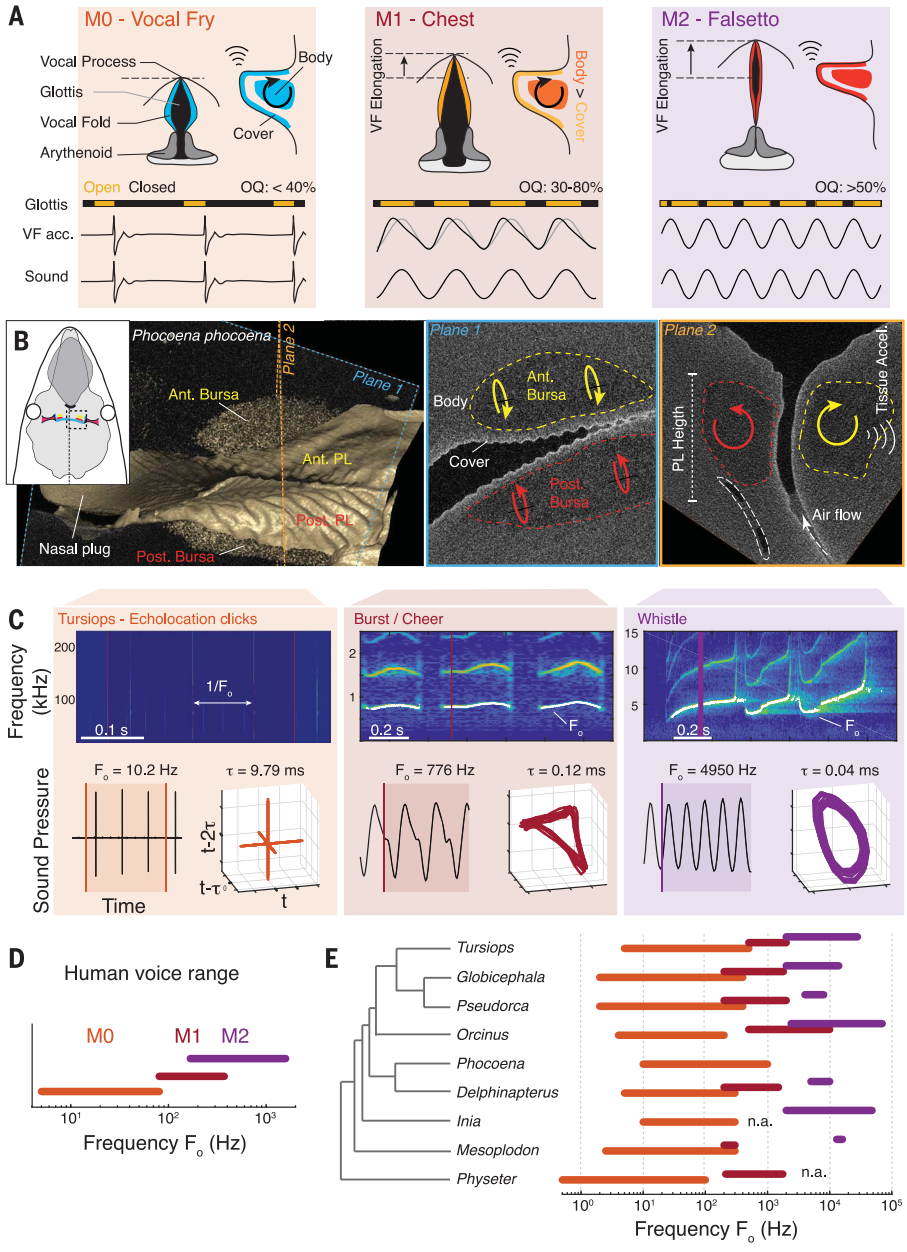


Fig. 3. The odontocete vocal repertoire and anatomy supports different vocal registers. (**A**) The human laryngeal vocal registers, M0 to M2. OQ, open quotient. (**B**) MicroCT scan of porpoise right PL. (**C**) Examples of bottlenose dolphin vocal repertoire with spectrogram (top), waveform (bottom left), and phase-space plot (bottom right) of click, burst, and whistle calls. (**D**) In humans, vocal registers M0 to M3 overlap in frequency ranges (table S1). (**E**) Frequency ranges of different vocalizations across odontocete clades support three vocal registers (table S1). N.a., no available data.

Discussion

We show that odontocetes have evolved an air-driven nasal sound source that is physically and functionally analogous to laryngeal and syringeal sound production in mammals and birds (*12*, *13*). The well-established myoelastic-aerodynamic theory for voiced sound production (*13*) provides a solid physiological basis to clasify odontocete vocal repertoires. Odontocetes use vocal registers to generate low-frequency, low-directionality, low–source level communication signals (*9*) and high-frequency, directional, high-power echolocation clicks. Vocal registers have been confirmed only in humans (*19*) and crows (*24*). They expand the f_o range and thus enhance vocal plasticity—a potential prerequisite for vocal learning. Bilateral specialization of phonic lip pairs may also aid in expanding the odontocete vocal range and allow simultaneous echolocation and communication (*9*). Acoustic source triangulation suggests that dolphins clicks are produced by right, and M1 and M2 vocalizations by the left phonic lips (*9*). As in songbirds, in which bilateral hemisyrinx tissues differentiate to produce different f_o ranges (*25*), odontocetes may use neural or anatomical specializations to exploit lateralized register vibrations.

We propose five advantages that drove stem odontocetes to evolve an air-driven nasal source to replace the ancestral laryngeal source.

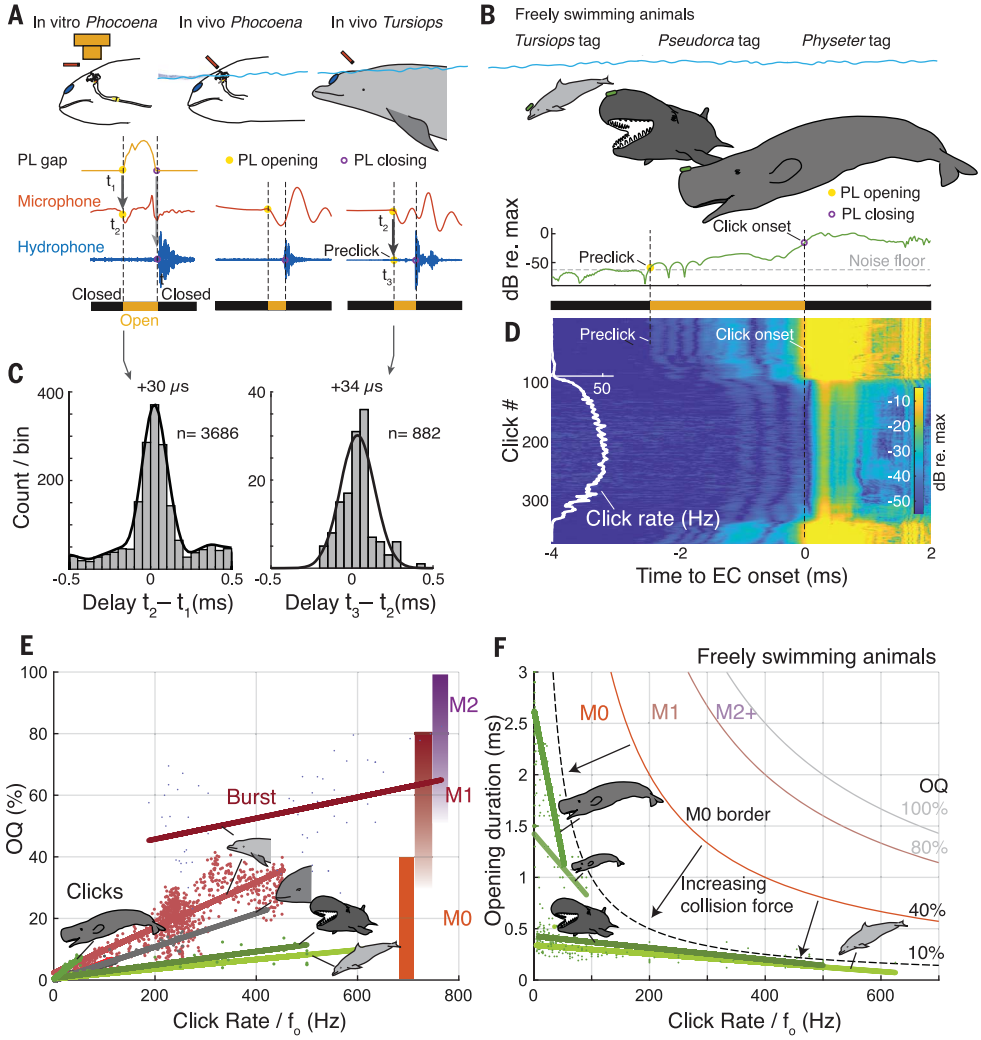


Fig. 4. The M0 register produces fast echolocation clicks in odontocetes. (**A**) PL opening and closing dynamics were reconstructed using acoustic signatures in restrained and (**B**) tagged, freely swimming odontocetes (*10*). (**C**) Delay between (left) PL opening and microphone pressure decrease and (right) the latter and hydrophone preclick onset. (**D**) Clicks time-aligned to onset in a freely hunting 35-ton sperm whale at 1800-m depth (fig S2). (**E**) OQ remains under 40% in trained and hunting animals, which coincides with the maximal OQ for M0 of 40% in humans. (**F**) Open-phase duration decreases with click rate so that OQ in free-swimming animals remains below 10%.

First, a nasal source freed the larynx from sound production, resulting in an effective valve that decoupled the lungs and nasal passages. This decoupling allows odontocetes 80-kPa driving pressures (*11*)—one order of magnitude above laryngeal driving pressures—to power the production of the highest biological source levels (*26*, *27*) without damaging lung tissues. Second, the nasal air volume is much smaller than that of the respiratory system, therefore allowing faster pressure control and air recycling. Third, the partially bone-lined nasal airspace provides mechanical resistance to hydrostatic compression and allows pressure gradient buildup over phonic lips at depths below ~100 m where the respiratory air volume matches the respiratory dead space (*28*) and lung air cannot be pressurized. Fourth, a sound source anterodorsal of the sound-reflecting skull generates a highly directional forward sound beam, acoustically isolated from their sensitive ears (*29*), which is not possible with laryngeally produced sounds. Fifth, by decoupling sound emission from their buccal cavity, odontocetes can catch and ingest food while echolocating. We suggest that because of anatomical adaptations to phonic lip closure, nasal-source evolution was primarily driven by selection on echolocation signals and only secondarily coopted for social communication.

Vocal fry (M0) vibration with ultrashort open times of 0.5 to 2.5 ms uncovers the biomechanical key to making prolonged click trains suitable for prey pursuit and capture. Because tissue acceleration directly translates to sound pressure in water, M0-register tissue acceleration generates short, broadband, powerful pressure transients in tissue and water that provide ideal long-range, high-frequency echolocation signals. Odontocetes thus exploited a feature of air-driven tissue vibration not advantageous in air. Furthermore, this air-driven system allows click repetition rates up to 500 clicks per second in the hunting phase prior to prey capture called the terminal buzz (*2*) (table S1), which is critical to active prey pursuit and capture. Our data show that the maximal buzz rate is set by the M0-to-M1 transition, which is determined by complex interplays of phonic lip size, tissue properties, posturing, and fluid-structure interaction (*21*, *22*). Just as in echolocating bats (*30*), the maximal buzz rate is thus facing a constraint related to sound production. Lastly, because of low OQ values, M0 is the most air-economical register (*31*). Odontocetes can make clicks with <50 μl of air per click (*8*). Such air economy allowed for echolocation at great depths, which opened the previously unexplored food niches of the deep ocean for exploitation.

REFERENCES AND NOTES

1. J. A. Goldbogen *et al.*, *Science* **366**, 1367–1372 (2019).
2. H. Vance *et al.*, *eLife* **10**, e68825 (2021).
3. T. Cranford *et al.*, *J. Exp. Mar. Biol. Ecol.* **407**, 81–96 (2011).
4. E. Chereskin *et al.*, *Curr. Biol.* **32**, 1657–1663.e4 (2022).
5. T. W. Cranford, M. Amundin, K. S. Norris, *J. Morphol.* **228**, 223–285 (1996).
6. S. Ridgeway *et al.*, in *Animal Sonar Systems*, R. G. Busnel, J. F. Fish, Eds. (Plenum, 1980), pp. 239–249.
7. K. Dormer, *J. Acoust. Soc. Am.* **65**, 229–239 (1979).
8. I. Foskolos, N. Aguilar de Soto, P. T. Madsen, M. Johnson, *Sci. Rep.* **9**, 15720 (2019).

9. P. T. Madsen, M. Lammers, D. Wisniewska, K. Beedholm, *J. Exp. Biol.* **216**, 4091–4102 (2013).
10. Materials and methods are available as supplementary materials.
11. M. Amundin, S. Andersen, *J. Exp. Biol.* **105**, 275–282 (1983).
12. C. P. Elemans *et al*., *Nat. Commun.* **6**, 8978 (2015).
13. J. G. Švec, H. K. Schutte, C. J. Chen, I. R. Titze, *J. Voice* **21**, 00055-2 (2021).
14. J. H. Lee *et al*., *Sci. Adv.* **4**, eaat9660 (2018).
15. J. J. Finneran *et al*., *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **204**, 271–283 (2018).
16. A. F. Mead *et al*., *eLife* **6**, e29425 (2017).
17. F. Ladich, H. Winkler, *J. Exp. Biol.* **220**, 2306–2317 (2017).
18. B. Jones, M. Zapetis, M. Samuelson, S. Ridgway, *Bioacoustics* **29**, 399–440 (2020).
19. B. Roubeau, N. Henrich, M. Castellengo, *J. Voice* **23**, 425–438 (2009).
20. H. Hollien, *J. Phonetics* **2**, 125–143 (1974).
21. I. T. Tokuda, J. Horácek, J. G. Svec, H. Herzel, *J. Acoust. Soc. Am.* **122**, 519–531 (2007).
22. Z. Zhang, *J. Acoust. Soc. Am.* **140**, 2614–2635 (2016).
23. S. Huggenberger, M. A. Rauschmann, T. J. Vogl, H. H. Oelschläger, *Anat. Rec. (Hoboken)* **292**, 902–920 (2009).
24. K. K. Jensen, B. G. Cooper, O. N. Larsen, F. Goller, *Proc. Biol. Sci.* **274**, 2703–2710 (2007).
25. R. Suthers, S. Zollinger, in *Neuroscience of Birdsong*, H. P. Zeigler, P. Marler, Eds. (Cambridge Univ. Press, 2008), pp. 78–98.
26. L. Jakobsen, J. Christensen-Dalsgaard, P. Juhl, C. P. H. Elemans, *Front. Ecol. Evol.* **9**, 657254 (2021).
27. B. Møhl, M. Wahlberg, P. T. Madsen, A. Heerfordt, A. Lund, *J. Acoust. Soc. Am.* **114**, 1143–1154 (2003).
28. A. Fahlman, M. J. Moore, D. Garcia-Parraga, *J. Exp. Biol.* **220**, 1761–1773 (2017).
29. F. H. Jensen, M. Johnson, M. Ladegaard, D. M. Wisniewska, P. T. Madsen, *Curr. Biol.* **28**, 3878–3885.e3 (2018).
30. C. P. Elemans, A. F. Mead, L. Jakobsen, J. M. Ratcliffe, *Science* **333**, 1885–1888 (2011).
31. R. E. McGlone, T. Shipp, *J. Speech Hear. Res.* **14**, 769–775 (1971).
32. N. D. Pyenson, *Curr. Biol.* **27**, R558–R564 (2017).

ACKNOWLEDGMENTS

We thank the Harderwijk staff—P. Bunskoek, N. van Elk, H. Goelema, S. Hearn, J. Mosterd, L. van Ooijen, and J. Wouters—for excellent support; H. Goelema and L. van Ooijen for training and scoping animals at Harderwijk; M. Ding for microCT scanning; M. Kieler, K. Beedholm, P. Tønnesen, and M. Ladegaard for discussions and analytical support; and W. T. Fitch, M. Johnson, J. Ratcliffe, M. Wahlberg, M. Amundin, and two anonymous reviewers for discussions and manuscript feedback. We dedicate this paper to the late Dr. Sam Ridgway. **Funding:** Danish research council grant 6108-00355A and Carlsberg Foundation grant CF16-0405 to P.T.M. and Carlsberg Foundation grant CF14-1096 and Novo Nordisk Foundation grant NNF20OC0063964 to C.P.H.E. **Author contributions:** Conceptualization: C.P.H.E. and P.T.M. Methodology: C.P.H.E. and P.T.M. In vivo and in vitro experimental design and analysis: C.P.H.E. Tag data acquisition and analysis: P.T.M. Porpoise head collection: U.S. Visualization: C.P.H.E. Funding acquisition: C.P.H.E. and P.T.M. Project administration: C.P.H.E. Writing – original draft: C.P.H.E. and P.T.M. Writing – review and editing: C.P.H.E., U.S., and P.T.M. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Code is available at the Zenodo Repository (*33*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adc9570
Materials and Methods
Figs. S1 to S5
Tables S1 and S2
References (*33*–*55*)
Movies S1 to S4

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

Submitted 12 May 2022; accepted 9 January 2023
10.1126/science.adc9570

PLANT SCIENCE

NLR immune receptor–nanobody fusions confer plant disease resistance

Jiorgos Kourelis†, Clemence Marchal†, Andres Posbeyikian, Adeline Harant, Sophien Kamoun*

Plant pathogens cause recurrent epidemics, threatening crop yield and global food security. Efforts to retool the plant immune system have been limited to modifying natural components and can be nullified by the emergence of new pathogen strains. Made-to-order synthetic plant immune receptors provide an opportunity to tailor resistance to pathogen genotypes present in the field. In this work, we show that plant nucleotide-binding, leucine-rich repeat immune receptors (NLRs) can be used as scaffolds for nanobody (single-domain antibody fragment) fusions that bind fluorescent proteins (FPs). These fusions trigger immune responses in the presence of the corresponding FP and confer resistance against plant viruses expressing FPs. Because nanobodies can be raised against most molecules, immune receptor–nanobody fusions have the potential to generate resistance against plant pathogens and pests delivering effectors inside host cells.

Plants lack an adaptive immune system and rely on innate immune receptors to detect invading pathogens. Efforts to retool the plant immune system to design new-to-nature biochemical activities have been largely limited to modification of natural components, for instance through receptor mutagenesis or domain shuffling (*1*–*8*). Although these approaches have yielded promising results, they often target a specific pathogen isolate and thus lack plasticity and adaptability to a wider range of pathogens and pests. Additionally, plant pathogens are notorious for rapidly evolving virulent races that can nullify new resistance specificities. Thus, there is a need for an adaptative system where resistance can be bioengineered as required to target the pathogen genotypes associated with plant disease outbreaks.

One class of immune proteins that could be optimal templates for receptor bioengineering is the subset of intracellular nucleotide-binding, leucine-rich repeat immune receptors (NLRs) that carry unconventional integrated domains (IDs) (*9*–*12*). These IDs are generally thought to mediate pathogen effector detection, either by directly binding to effectors or by acting as a substrate for their enzymatic activity. This activity is subsequently translated into an immune response (*13*–*17*). Often these ID-containing NLRs (NLR-IDs) are genetically linked to conventional NLRs that are required for immune activation after effector detection (*9*, *18*). Pik-1 and Pik-2 are such an NLR receptor pair from rice carrying an N-terminal coiled-coil (CC) domain (*19*). Pik-1 carries an integrated heavy metal–associated (HMA) domain between its CC and the central NB-ARC (nucleotide-binding domain shared with APAF-1, various R proteins, and CED-4) domains that directly binds AVR-Pik effector proteins (avirulence factors) secreted by the blast fungus, *Magnaporthe oryzae* (*13*, *16*, *20*–*24*). AVR-Pik binding to the Pik-1 HMA domain results in Pik-2–dependent immunity (*16*, *19*). The integrated HMA domain of Pik-1 can be mutated or swapped for similar HMA domains to confer recognition of different AVR-Pik alleles (*5*, *25*, *26*). The Pik NLR gene pair has a variety of alleles (such as Pikp, Pikm, Pikh, Piks, and Pik*) that provide immunity to blast isolates expressing different AVR-Pik variants (*27*).

What would be the ultimate ID for engineering made-to-order plant immune receptors? Given that animal adaptive immunity has the capacity to generate antibodies against virtually any antigen that it is exposed to, we reasoned that harnessing antibodies for plant immunity would potentially enable building receptors that respond to a wide range of plant pathogen molecules. We focused on the minimal antigen-binding fragment of single-domain heavy-chain antibodies (known as VHHs or nanobodies) of camelid mammals (*28*–*31*) because they are small, soluble 10- to 15-kDa domains, which tend to correctly fold intracellularly and have many useful properties for biotechnological applications. To test our idea, we generated orthogonal Pik-1 sensors in which the integrated HMA domain is swapped with nanobodies that bind either green fluorescent protein (GFP) or mCherry (*32*–*34*) (Fig. 1, A and B, and table S1). We hypothesized that the engineered versions of Pik-1 would trigger immunity in the presence of GFP or mCherry.

Pikobodies are functional NLR-nanobody fusions

Mutations in the Pik-1 HMA domain often lead to autoimmune activities in the absence of a ligand, and like the immune signaling in response to effector recognition, this activity is dependent on the presence of Pik-2 (*23*, *25*, *26*, *35*). Hence, we first tested whether the Pikm-1–nanobody fusions induce autoimmunity in the presence of Pik-2. Of the 11 tested Pikm-1–nanobody fusions, six did not exhibit autoimmunity when expressed with Pikm-2 in leaves of the model plant *Nicotiana benthamiana* (fig. S1 and table S2), which indicated that they can be used for follow-up gain-of-function assays. Next, we co-expressed 10 Pikm-1–nanobody fusions with GFP or mCherry. Among these, four produced a hypersensitive cell death response (HR, immune response readout) specifically when expressed with their matching fluorescent proteins (FPs) (Enhancer, LaG-16, LaM-4, and LaM-8) (Fig. 1C, fig. S1, and table S2). The response levels were similar to those obtained with a natural combination of Pikm and a blast fungus effector (Fig. 1C and fig. S1). Additionally, a further three fusions that displayed weak autoimmunity gave a stronger HR only when combined with their matching FPs (LaG-24, LaM-2, and LaM-6) (Fig. 1C and fig. S1). This indicates that the Pikm-1–nanobody fusions are functional and can be endowed with new-to-nature activities. We coined the term Pikobody for the combination of Pikm-2 with the engineered Pikm-1–nanobody fusions (Fig. 1).

We reasoned that nanobody aggregation or misfolding upon intracellular expression (*32*) could explain the observed autoactivity (Fig. 1C and fig. S1). We introduced previously described stabilizing nanobody mutations (*32*) in LaG-24, LaM-2, LaM-3, and LaM-6 and found that they abolished Pikobody autoactivity (fig. S2). Three of these Pikobodies carrying the stabilized mutants of LaG-24, LaM-3, and LaM-6 retained the capacity to trigger HR in the presence of the matching FP, whereas LaM-2 did not (fig. S2). This shows that Pikobody autoactivity can be abolished by engineering the core structural features of the nanobodies.

Given that Enhancer and LaM-4 are widely used nanobodies recognizing GFP and mCherry, respectively, we selected Pikobody^{Enhancer} (consisting of Pikm-2 together with the Pikm-1^{Enhancer}–nanobody fusion) and Pikobody$^{LaM\text{-}4}$ (consisting of Pikm-2 together with the Pikm-1$^{LaM\text{-}4}$–nanobody fusion) to further confirm our results. Some pathogen effectors may not be recognized by Pikobodies, for example by having a subcellular localization that precludes recognition. We noted that Pikobody^{Enhancer} and Pikobody$^{LaM\text{-}4}$ specifically responded to three pathogen effectors only when they were tagged with the matching GFP or enhanced GFP (EGFP) or the matching mCherry or mRFP1 (fig. S3 and table S2). This further confirmed that the Pikobodies are functional FP sensors that detect FPs even when they are fused to pathogen effector proteins.

We investigated the extent to which Pikobodies function through similar mechanisms

The Sainsbury Laboratory, University of East Anglia, Norwich Research Park, Norwich NR4 7UH, UK.
*Corresponding author. Email: sophien.kamoun@tsl.ac.uk
†These authors contributed equally to this work.

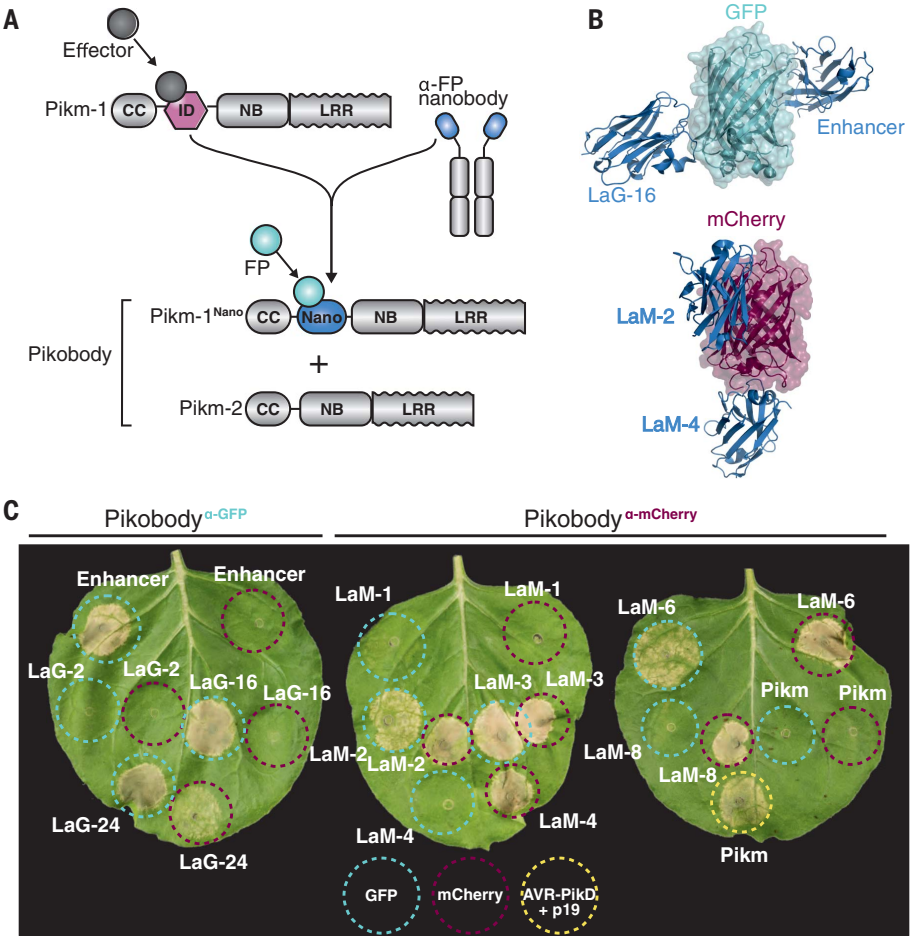


Fig. 1. NLR immune receptor–nanobody fusions trigger a HR in the presence of the corresponding FP antigen. (**A**) Bioengineering of a FP-activated NLR sensor. The integrated HMA domain of the NLR Pikm-1, which is involved in pathogen effector recognition by direct binding, was swapped with nanobodies binding either GFP or mCherry. We coined the term Pikobody for the combination of Pikm-2 with Pikm-1–nanobody fusions. CC, coiled-coil domain; NB, nucleotide-binding domain; LRR, leucine-rich repeat; Nano, nanobody. (**B**) Structures of GFP (cyan) or mCherry (magenta) with the GFP-binding nanobodies Enhancer [Protein Data Bank (PDB) ID: 3K1K] (*33*) and LaG-16 (PDB: 6LR7) (*51*) or the mCherry-binding nanobodies LaM-2 (PDB: 6IR2) and LaM-4 (PDB: 6IR1) (*52*), respectively (dark blue). (**C**) Screen for GFP or mCherry recognition by bioengineered Pikobody$^{\alpha\text{-GFP}}$ or Pikobody$^{\alpha\text{-mCherry}}$ resulting in induction of a HR upon coexpression in *N. benthamiana*. Representative *N. benthamiana* leaves were infiltrated with the indicated constructs and photographed 5 days after infiltration (see fig. S1 for quantification). Cyan and magenta dashed circles indicate GFP or mCherry co-infiltration, respectively. The Pikm pair (Pikm-1 and Pikm-2) co-infiltrated with AVR-PikD, and p19 was used as a positive control for HR (yellow).

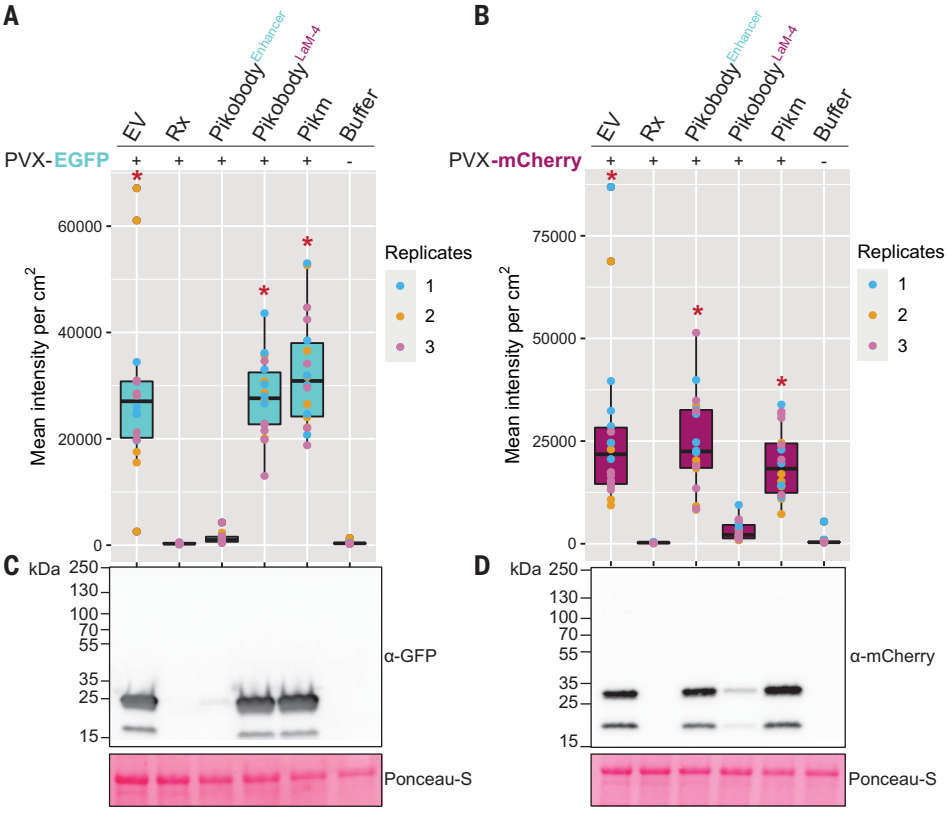


Fig. 2. Pikobodies confer resistance against PVX variants expressing matching FPs. (**A** and **B**) Specific reduction in fluorescence intensity of PVX-expressed EGFP or mCherry in the presence of Pikobody$^{\text{Enhancer}}$ or Pikobody$^{\text{LaM-4}}$, respectively. GFP (A) or mCherry (B) mean fluorescence intensity per square centimeter measured in *N. benthamiana* leaves 4 days after infiltration and used as a proxy for PVX viral load. Boxplots summarize results of three independent replicates with six internal replicates. Asterisks show significant differences between buffer only (no PVX added) and tested constructs in the presence of PVX-GFP (A) or PVX-mCherry (B) (Dunnett's test, $P < 0.001$). The PVX coat protein recognizing resistance protein Rx was used as a positive control for PVX resistance. (**C** and **D**) Specific reduction of PVX-expressed GFP (C) or mCherry (D) accumulation as a proxy to evaluate viral load in the presence of Pikobody$^{\text{Enhancer}}$ or Pikobody$^{\text{LaM-4}}$, respectively. Total protein was extracted 4 days after inoculation with PVX variants in the presence of the tested constructs and probed with the appropriate antibodies. Ponceau S staining shows equal protein loading across samples.

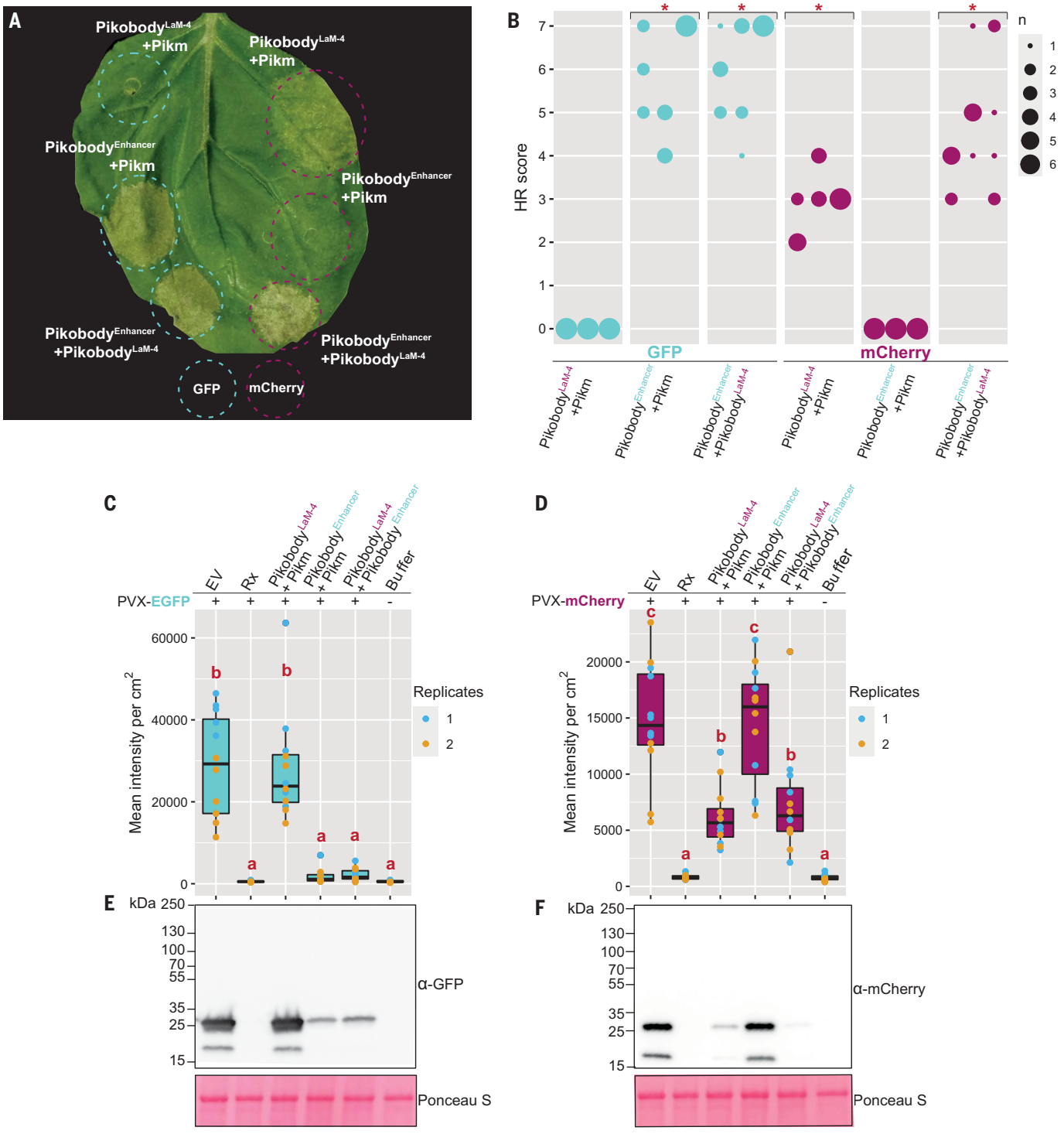


Fig. 3. Stacked Pikobodies result in additive immune recognition and disease resistance. (**A**) Pikobody stacking results in additive immune recognition. A representative *N. benthamiana* leaf infiltrated with indicated constructs is shown. Cyan or magenta dashed lines indicate GFP or mCherry coexpression, respectively. Leaves were photographed 4 days after infiltration. (**B**) HR quantification visualized as a dot plot, where the size of a dot is proportional to the number of samples with the same score (*n*) within the same replicate (1 to 3). The experiment was repeated three times with six internal replicates. Asterisks indicate statistically significant differences as compared with the Pikobody^{Enhancer}+Pikm+mCherry control, as determined by the besthr R package (*53*). (**C** and **D**) Specific reduction in fluorescence intensity of PVX-expressed GFP (C) and mCherry (D) in the presence of stacked Pikobody^{Enhancer} and Pikobody^{LaM-4}. Mean fluorescence intensity per square centimeter was measured in *N. benthamiana* leaves 4 days after infiltration and used as a proxy for PVX viral load. Boxplots summarize the results of three independent replicates with six internal replicates. Letters depict significant differences between treatments as determined by analysis of variance (ANOVA) followed by Tukey's honestly significant difference (HSD) test ($P < 0.05$). (**E** and **F**) Specific reduction of PVX-expressed GFP (E) or mCherry (F) accumulation as a proxy to evaluate viral load in the presence of stacked Pikobody^{Enhancer} or Pikobody^{LaM-4}. Total protein was extracted 4 days after inoculation with PVX variants in the presence of the tested constructs and probed with the appropriate antibodies. Ponceau S staining shows equal protein loading across samples.

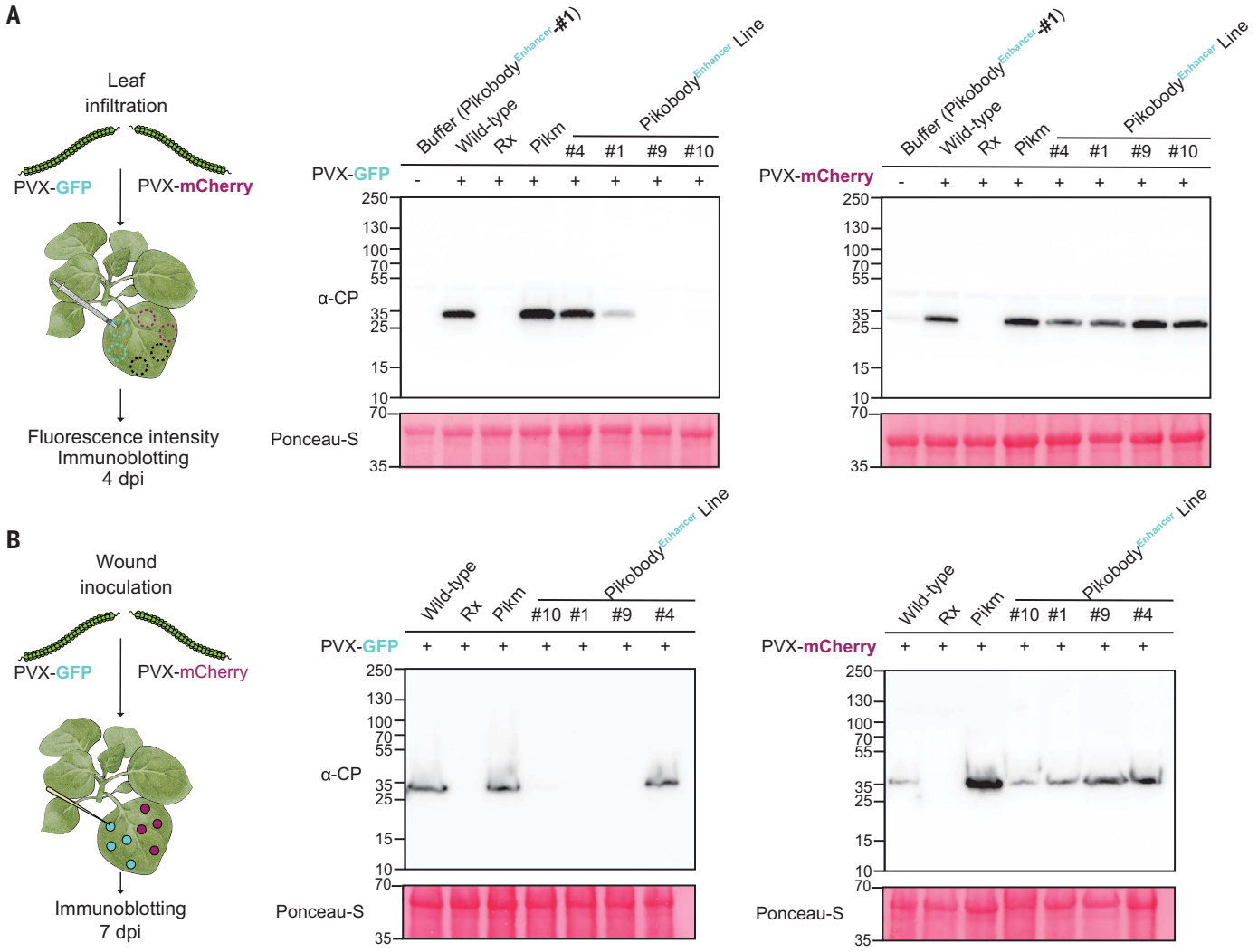


Fig. 4. Transgenic *N. benthamiana* Pikobody^{Enhancer} lines confer specific resistance to PVX-GFP. (**A**) Pikobody^{Enhancer} line 9 displays specific resistance to PVX-GFP—but not to PVX-mCherry—delivered by leaf agroinfiltration to an extent similar to that displayed by Rx. Coat protein (CP) accumulation was used to directly measure PVX viral load. For the immunoblot analysis, total protein was extracted 4 days after inoculation of the stated lines with PVX-GFP or PVX-mCherry and probed with the corresponding antibodies. Buffer only (no added PVX) was used as a negative control. Ponceau S staining shows equal protein loading across samples. dpi, days postinfiltration. (**B**) Pikobody^{Enhancer} lines 1 and 9 display specific resistance to PVX-GFP—but not to PVX-mCherry—delivered by wounding (toothpick inoculation of agrobacterium) to an extent similar to that of Rx. Immunoblot analysis is as in (A).

as the wild-type Pik pair (*36*) and other CC-NLRs. The conserved P-loop motif within the NB-ARC domain of CC-NLRs is required for the adenosine 5′-diphosphate (ADP)–adenosine 5′-triphosphate (ATP) switch that enables oligomerization into resistosome complexes (*37*, *38*). Pikobody^{K217R,Enhancer} and Pikobody$^{K217R,LaM-4}$ with a P-loop dead mutation in Pikm-2 (Pikm-2^{K217R}) failed to produce a HR to their corresponding FP even though the Pikm-1 and Pikm-2 proteins accumulated to similar levels as the wild-type immune receptors (fig. S4 and table S2). We conclude that the P-loop motif of Pikm-2 is required for Pikobody activity and that the Pikobody system probably functions through the established mechanistic model of NLRs (*39*).

Transient Pikobody expression confers immunity to PVX

Can Pikobodies produce a functional immune response that is effective against a pathogen? We used recombinant *Potato virus X* (PVX) (*40*) expressing either GFP or mCherry to assay the ability of Pikobodies to reduce viral load (table S1). These PVX variants express FPs from a duplicated coat protein subpromoter in the virus genome. We used fluorescence intensity and immunodetection of GFP or mCherry accumulation as proxy for viral load in leaf samples (Fig. 2 and table S3). Both Pikobody^{Enhancer} and Pikobody^{LaM-4} specifically reduced fluorescence intensity of PVX-expressed GFP or mCherry, respectively, to an extent comparable to that of Rx, an NLR known to confer immunity against PVX (*41*) (Fig. 2, A and B). This reduction of fluorescence intensity correlates with reduced accumulation of virus-expressed GFP or mCherry as compared with the empty vector control or wild-type Pikm (Fig. 2, C and D). We did, however, observe a faint signal corresponding to GFP or mCherry in the samples with PVX-GFP or PVX-mCherry and Pikobody^{Enhancer} or Pikobody^{LaM-4}, respectively, as compared with no detectable FP bands in the samples with Rx (Fig. 2, C and D).

To independently confirm these results, we tested two additional PVX variants expressing GFP from different virus genome locations (*42*, *43*) (tables S1 and S3 and fig. S5). Pikobody^{Enhancer} but not Pikobody^{LaM-4} reduced GFP fluorescence intensity and protein

accumulation when challenged with a PVX variant with the GFP sequence inserted between the triple gene block and coat protein in the virus genome (fig. S5, A and B). Furthermore, we observed reduced accumulation of virus-expressed GFP in the presence of Pikobody^{Enhancer} when challenged with a PVX variant carrying an in-frame N-terminal GFP fusion to the viral coat protein separated by a foot-and-mouth disease virus 2A self-cleaving peptide, despite consistently not seeing a significant change in fluorescence intensity in the presence of Pikobody^{Enhancer} (fig. S5, C and D). GFP fluorescence is known to be enhanced in the presence of certain nanobodies (*33*), and perhaps the fluorescent properties of the population of GFP that remains fused to the coat protein are enhanced under these conditions (*43*) (fig. S5D). Additionally, at late stages of infection Pikobody^{Enhancer} but not Pikobody$^{LaM\text{-}4}$ caused a visible HR with both variants of PVX-GFP (fig. S5, E and F).

We further compared the virus resistance capacity of Pikobody^{Enhancer} with the disease resistance protein Rx using PVX variants that contain mutations in the coat protein that evade Rx-mediated immunity (*44*). Unlike Rx, Pikobody^{Enhancer} conferred resistance against the Rx-evading variant PVX-GFP::CP^{T122K,M128R} (fig. S6 and table S3). We conclude that Pikobody^{Enhancer} can provide resistance against multiple PVX-GFP variants, including a variant that evades Rx.

Pikobody stacking results in additive recognition capacities

The addition of more than one immune receptor in a plant variety—a plant breeding strategy known as *R* gene stacking—can maximize resistance durability in the field by delaying the emergence of virulent pathogen races (*45*–*47*). However, coexpression of plant immune receptors can lead to autoimmunity (*48*, *49*) or suppression of recognition (*50*). We investigated whether Pikobodies with different FP specificities are compatible with each other (Fig. 3). We first determined that the coexpression of different Pikobodies does not result in autoimmunity or affect Pikobody accumulation (fig. S7 and table S2). Coexpression of Pikobody^{Enhancer} or Pikobody$^{LaM\text{-}4}$ and the wild-type Pikm pair triggered a HR only in the presence of the matching FP, whereas coexpression of Pikobody^{Enhancer} and Pikobody$^{LaM\text{-}4}$ produced a HR in the presence of both GFP and mCherry (Fig. 3, A and B). Similarly, coexpression of Pikobody^{Enhancer} and Pikobody$^{LaM\text{-}4}$ markedly reduced fluorescence intensity and protein levels of both GFP and mCherry produced by PVX-FPs (Fig. 3, C to F). At late stages of infection, the combination of Pikobody^{Enhancer} and Pikobody$^{LaM\text{-}4}$ also resulted in a HR in response to either PVX-GFP or PVX-mCherry (fig. S8). We conclude that Pikobody stacking can expand the recognition and response profile of these immune receptors without necessarily resulting in autoimmunity.

Transgenic plants expressing Pikobodies are resistant to PVX

We challenged our findings that Pikobodies confer virus resistance using stable *N. benthamiana* transgenic lines expressing Pikobody^{Enhancer} (Fig. 4). Among four lines transformed with Pikobody^{Enhancer}, three specifically respond to GFP, accumulate Pikm-2 and Pikm-1^{Enhancer} proteins, and reduce virus load of PVX-GFP but not of PVX-mCherry, as estimated by the accumulation of FPs and PVX coat protein (lines 1, 9, and 10) (fig. S9 and tables S4 and S5). A fourth transgenic that does not respond to GFP and only accumulates Pikm-2 protein and not Pikm-1^{Enhancer} protein served as a negative control (line 4) (Fig. 4 and fig. S10). The level of PVX resistance in one of the Pikobody^{Enhancer} transgenic lines (line 9) was similar to that of Rx, with no detectable levels of GFP and coat protein (Fig. 4 and fig. S10). Notably, the Pikobody^{Enhancer} transgenic lines conferred resistance to PVX regardless of the agroinfection method used to deliver the virus (leaf infiltration and wounding inoculation) (Fig. 4 and figs. S10 and S11). We conclude that transgenic Pikobody lines can confer specific resistance to PVX to a similar extent as the natural resistance gene Rx.

Discussion and conclusion

We built on our growing understanding of the evolution and function of the Pik pair of NLRs (*5*, *21*–*23*, *36*) to use Pik-1 as a chassis for VHH nanobody fusions to bioengineer functional disease resistance genes with new-to-nature functionalities. This strategy for bioengineering synthetic immune receptors differs from earlier approaches, which were based on the modification of endogenous sequences and domains. The Pikobody system provides a method to functionally transfer components of the metazoan immune system to plants. Given that Pikobodies rely on NLR-mediated immunity, this system shares the same limitations as other approaches that leverage this immune receptor family. For example, pathogen proteins will need to be translocated inside the plant cell during the right phase of pathogen infection to be recognized by Pikobodies. Nevertheless, given that nanobodies can be readily generated to bind virtually any antigen, Pikobodies have the potential to produce made-to-order resistance genes against any pathogen or pest that delivers effectors inside host plant cells (fig. S12).

REFERENCES AND NOTES

1. S. H. Kim, D. Qi, T. Ashfield, M. Helm, R. W. Innes, *Science* **351**, 684–687 (2016).
2. S. E. Pottinger *et al.*, *Mol. Plant Microbe Interact.* **33**, 932–944 (2020).
3. M. E. Segretin *et al.*, *Mol. Plant Microbe Interact.* **27**, 624–637 (2014).
4. A. Giannakopoulou *et al.*, *Mol. Plant Microbe Interact.* **28**, 1316–1329 (2015).
5. J. C. De la Concepcion *et al.*, *eLife* **8**, e47713 (2019).
6. S. Césari *et al.*, *Nat. Commun.* **13**, 1524 (2022).
7. S. Wang, W. Huang, Z. Duxbury, S. A. Hogenhout, J. D. G. Jones, bioRxiv 2021.09.06.459143 [Preprint] (2021). https://doi.org/10.1101/2021.09.06.459143.
8. J. L. Dangl, D. M. Horvath, B. J. Staskawicz, *Science* **341**, 746–751 (2013).
9. S. Césari, M. Bernoux, P. Moncuquet, T. Kroj, P. N. Dodds, *Front. Plant Sci.* **5**, 606 (2014).
10. C.-H. Wu, K. V. Krasileva, M. J. Banfield, R. Terauchi, S. Kamoun, *Front. Plant Sci.* **6**, 134 (2015).
11. T. Kroj, E. Chanclud, C. Michel-Romiti, X. Grand, J.-B. Morel, *New Phytol.* **210**, 618–626 (2016).
12. P. F. Sarris, V. Cevik, G. Dagdas, J. D. G. Jones, K. V. Krasileva, *BMC Biol.* **14**, 8 (2016).
13. S. Césari *et al.*, *Plant Cell* **25**, 1463–1481 (2013).
14. P. F. Sarris *et al.*, *Cell* **161**, 1089–1100 (2015).
15. C. Le Roux *et al.*, *Cell* **161**, 1074–1088 (2015).
16. A. Maqbool *et al.*, *eLife* **4**, e08709 (2015).
17. L. Guo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 11637–11642 (2018).
18. T. K. Eitas, J. L. Dangl, *Curr. Opin. Plant Biol.* **13**, 472–477 (2010).
19. I. Ashikawa *et al.*, *Genetics* **180**, 2267–2276 (2008).
20. H. Kanzaki *et al.*, *Plant J.* **72**, 894–907 (2012).
21. J. C. De la Concepcion *et al.*, *Nat. Plants* **4**, 576–585 (2018).
22. J. C. De la Concepcion *et al.*, *PLOS Pathog.* **17**, e1009368 (2021).
23. A. Białas *et al.*, *eLife* **10**, e66961 (2021).
24. A. Białas *et al.*, *Mol. Plant Microbe Interact.* **31**, 34–45 (2018).
25. J. H. R. Maidment *et al.*, bioRxiv 2022.06.14.496076 [Preprint] (2022). https://doi.org/10.1101/2022.06.14.496076.
26. A. R. Bentham *et al.*, bioRxiv 2022.10.10.511592 [Preprint] (2022). https://doi.org/10.1101/2022.10.10.511592.
27. K. Yoshida *et al.*, *Plant Cell* **21**, 1573–1591 (2009).
28. C. Hamers-Casterman *et al.*, *Nature* **363**, 446–448 (1993).
29. A. S. Greenberg *et al.*, *Nature* **374**, 168–173 (1995).
30. S. Muyldermans, *Annu. Rev. Biochem.* **82**, 775–797 (2013).
31. D. Könning *et al.*, *Curr. Opin. Struct. Biol.* **45**, 10–16 (2017).
32. J. G. Dingus, J. C. Y. Tang, R. Amamoto, G. K. Wallick, C. L. Cepko, *eLife* **11**, e68253 (2022).
33. A. Kirchhofer *et al.*, *Nat. Struct. Mol. Biol.* **17**, 133–138 (2010).
34. P. C. Fridy *et al.*, *Nat. Methods* **11**, 1253–1260 (2014).
35. J. C. De la Concepcion *et al.*, *eLife* **10**, e71662 (2021).
36. R. Zdrzałek, S. Kamoun, R. Terauchi, H. Saitoh, M. J. Banfield, *PLOS ONE* **15**, e0238616 (2020).
37. J. Wang *et al.*, *Science* **364**, eaav5868 (2019).
38. S. P. Dinesh-Kumar, W.-H. Tham, B. J. Baker, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14789–14794 (2000).
39. F. L. W. Takken, W. I. L. Tameling, *Science* **324**, 744–746 (2009).
40. S. Marillonnet, C. Engler, V. Klimyuk, Y. Gleba, Potexvirus-derived replicon, Patent WO/2008/028661 (2008); https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2008028661.
41. A. Bendahmane, K. Kanyuka, D. C. Baulcombe, *Plant Cell* **11**, 781–792 (1999).
42. R. Lu *et al.*, *EMBO J.* **22**, 5690–5699 (2003).
43. S. S. Cruz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6286–6290 (1996).
44. M. G. Goulden, B. A. Köhm, S. Santa Cruz, T. A. Kavanagh, D. C. Baulcombe, *Virology* **197**, 293–302 (1993).
45. M. Luo *et al.*, *Nat. Biotechnol.* **39**, 561–566 (2021).
46. S. Zhu, Y. Li, J. H. Vossen, R. G. F. Visser, E. Jacobsen, *Transgenic Res.* **21**, 89–99 (2012).
47. M. Ghislain *et al.*, *Plant Biotechnol. J.* **17**, 1119–1129 (2019).
48. E. Chae *et al.*, *Cell* **159**, 1341–1351 (2014).
49. D. T. N. Tran *et al.*, *Curr. Biol.* **27**, 1148–1160 (2017).
50. S. Hurni *et al.*, *Plant J.* **79**, 904–913 (2014).
51. Z. Zhang, Y. Wang, Y. Ding, M. Hattori, *Sci. Rep.* **10**, 6239 (2020).
52. Z. Wang *et al.*, *Protein Sci.* **30**, 2298–2309 (2021).
53. D. MacLean, besthr - Generating Bootstrap Estimation Distributions of HR Data, Github (2020); https://github.com/TeamMacLean/besthr.

ACKNOWLEDGMENTS

We thank the Tissue Culture and Transformation Team at the Sainsbury Laboratory for generating stable *N. benthamiana* transgenic lines containing the Pikobody system. We thank H. Pai for drawing the llama and the *N. benthamiana* plants in figs. S10 to S12, A. Białas for helpful comments on the figures, A. Williams for useful comments on the text, and P. Robinson for photography. We also thank our long-standing collaborators M. Banfield, N. Talbot, R. Terauchi, and other members of the BLASTOFF

community for the many useful discussions and suggestions. **Funding:** The authors received funding from the Gatsby Charitable Foundation (C.M., A.P., and S.K.), Biotechnology and Biological Sciences Research Council (BBSRC) BB/P012574 (Plant Health ISP) (S.K.), European Research Council (ERC) 743165 (A.H. and S.K.), and BASF Plant Science (J.K. and S.K.). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. **Author contributions:** Conceptualization: J.K., C.M., and S.K. Methodology: J.K. and C.M. Formal analysis: J.K., C.M., and A.P. Investigation: J.K., C.M., A.P., and A.H. Resources: J.K. Writing – original draft: J.K., C.M., and S.K. Writing – review & editing: J.K., C.M., A.P., A.H., and S.K. Visualization: J.K. and C.M. Supervision: S.K. Project administration: S.K. Funding acquisition: S.K. **Competing interests:** J.K., C.M., and S.K. receive funding from industry on NLR biology and have filed a patent on receptor-nanobody fusions (European patent application no. 21386064.6). The authors declare no other competing interests. **Data and materials availability:** Pikobody constructs generated for this study will be subjected to material transfer agreements (MTAs) and made available upon request. All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn4116
Materials and Methods
Figs. S1 to S12
Tables S1 to S6
References (*54*–*64*)
MDAR Reproducibility Checklist

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

Submitted 25 November 2021; resubmitted 9 December 2022
Accepted 1 February 2023
10.1126/science.abn4116

RESEARCH ARTICLE SUMMARY

DRUG DISCOVERY

Abiotic peptides as carriers of information for the encoding of small-molecule library synthesis

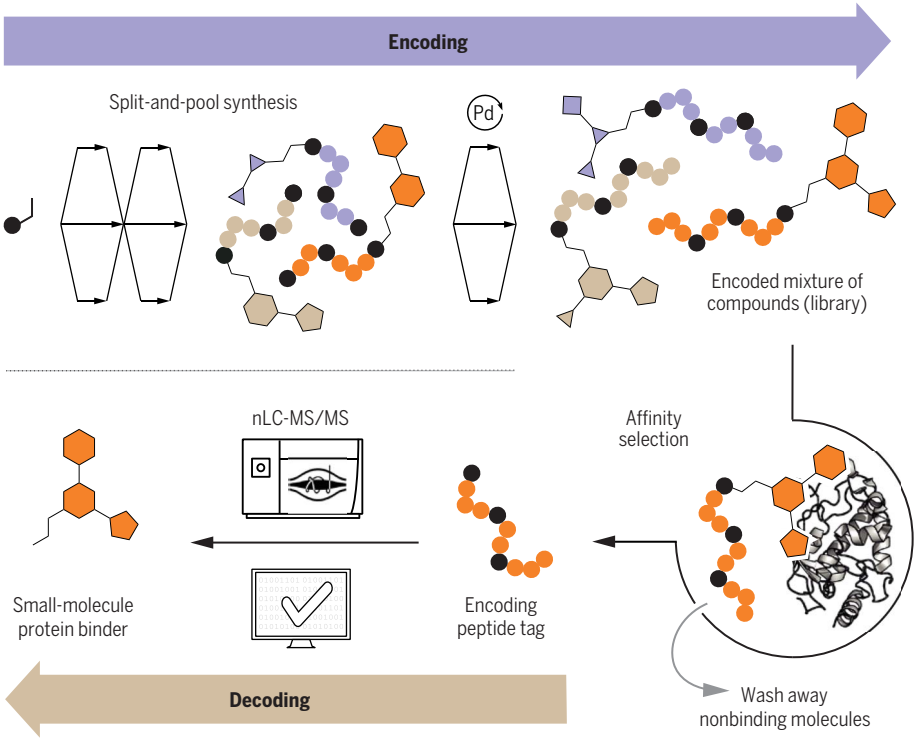
Simon L. Rössler†, Nathalie M. Grob†, Stephen L. Buchwald*, Bradley L. Pentelute*

INTRODUCTION: The discovery of therapeutics and biochemical probes hinges on the ability to identify molecules that interact with biological targets of interest. Technologies such as DNA-encoded libraries have transformed the process of drug discovery by enabling the rapid synthesis of vast collections of molecules, each encoded with a unique appendant DNA tag, and their subsequent screening in affinity selection experiments. However, the use of DNA to encode information confers synthetic limitations owing to the incompatibility of oligonucleotides with various chemical reaction conditions that result in the loss of stored information. As a consequence, chemical transformations developed for DNA-encoded library synthesis have to be optimized for oligonucleotide compatibility rather than reaction efficiency and scope. Given the vast potential of encoded library technologies in drug discovery, complementary platforms addressing the limitations of DNA encoding by leveraging carriers of information with higher stability and versatility are desirable.

RATIONALE: Molecular encoding can in principle be achieved in any polymer with at least two distinguishable monomers. Peptides constitute a biopolymer that is routinely sequenced for elucidation of protein identity through the use of tandem mass spectrometry. Accordingly, synthetic peptides can serve as carriers of information decoded through determination of their amino acid sequence. We hypothesized that the excellent chemical stability of peptides and their compatibility with a broad range of reaction conditions would render them particularly suited for the encoding of small-molecule synthesis. Peptides, which can be efficiently prepared by solid-phase synthesis, offer potential as carriers of information for a next-generation discovery platform with combinatorial libraries of small molecules.

RESULTS: We describe the design of abiotic peptides as carriers of information for the encoding of small-molecule synthesis. Therein, the identity of a small molecule was stored in an appendant peptide that is elongated in accordance with the synthetic elaboration of the small molecule. The encoding peptide featured a hexadecimal encoding alphabet of nonisobaric amino acids, resulting in high information density and chemical stability. The sequence of the encoding peptide was optimized through the systematic inclusion of selected amino acids to fine-tune polarity and ease of sequencing, resulting in high-fidelity decoding by tandem mass spectrometry. The chemical stability of the peptide tag enabled synthetic versatility for small-molecule transformations, including acidic conditions or transition metal catalysis with reported incompatibility with DNA tags. This broad compatibility allowed the implementation of palladium-mediated cross-coupling reactions characterized by a diverse scope and high reaction efficiency. The encoding of small molecule synthesis in peptides was leveraged for the split-and-pool synthesis of combinatorial libraries called peptide-encoded libraries (PELs) characterized by high purity. PELs featuring tens of thousands of drug-like small molecules resulting from optimized palladium-mediated C–C and C–N cross-coupling reactions were used in affinity selections against oncogenic proteins. The peptide sequences of enriched conjugates were decoded, and the corresponding small molecules were rapidly prepared by solid-phase synthesis and subsequently confirmed to exhibit affinity for their target protein. The PEL discovery platform is characterized by high efficiency and has afforded diverse, previously unknown small-molecule binders for the target proteins.

CONCLUSION: The results demonstrate that abiotic peptides can be used to encode and decode information to discover small molecules with affinity to proteins of interest. The PEL discovery platform establishes a starting point for the next generation of encoded library technology with broad implications for therapeutics discovery and biochemical research. ■



Peptide-encoded libraries of small molecules enable the discovery of hit compounds with affinity for protein targets of interest. Split-and-pool synthesis of small molecules generates libraries encoded by unique peptide tags. The peptide-encoded libraries are used in affinity selections with targets of interest. After releasing the encoding tag, the information is decoded by tandem mass spectrometry to identify the hit molecule. nLC-MS/MS, nanoscale liquid chromatography–tandem mass spectrometry.

The list of author affiliations is available in the full article.
*Corresponding author. Email: sbuchwal@mit.edu (S.L.B.); blp@mit.edu (B.L.P.)
†These authors contributed equally to this work.

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adf1354

DRUG DISCOVERY

Abiotic peptides as carriers of information for the encoding of small-molecule library synthesis

Simon L. Rössler[1]†, Nathalie M. Grob[1]†, Stephen L. Buchwald[1]*, Bradley L. Pentelute[1,2,3,4]*

Encoding small-molecule information in DNA has been leveraged to accelerate the discovery of ligands for therapeutic targets such as proteins. However, oligonucleotide-based encoding is hampered by inherent limitations of information stability and density. In this study, we establish abiotic peptides for next-generation information storage and apply them for the encoding of diverse small-molecule synthesis. The chemical stability of the peptide-based tag allows the use of palladium-mediated reactions to efficiently synthesize peptide-encoded libraries (PELs) with broad chemical diversity and high purity. We demonstrate the successful de novo discovery of small-molecule protein ligands from PELs by affinity selection against carbonic anhydrase IX and the oncogenic protein targets BRD4(1) and MDM2. Collectively, this work establishes abiotic peptides as carriers of information for the encoding of small-molecule synthesis, leveraged herein for the discovery of protein ligands.

The efficient storage of information in biological macromolecules forms the fundamental basis of cellular life. DNA, nature's carrier of information, offers unparalleled efficiency for the encoding and decoding of information (*1*). Endeavors to extend this efficiency to non-natural systems have afforded promising advances in information storage technologies and drug discovery platforms (*2*, *3*). Most prominently, DNA-encoded libraries (DELs) enable identification of potential ligands for protein targets through the exploration of chemical space on an enormous scale (*4*, *5*). However, DNA-based technologies face limitations inherent to their oligonucleotide building blocks. Depurination or strand cleavage with concomitant loss of encoded information may occur under nonphysiological conditions, particularly in reaction environments required for chemical synthesis that involve metals, acids, oxidants, and radical species (*6*). Additionally, DNA-based discovery platforms may be unsuited to target DNA binding proteins of therapeutic relevance such as transcription factors. The widespread interest in DELs over the past decades has afforded several pioneering remedies that enable broader applications (*7*–*13*), however, inherent limitations associated with the use of oligonucleotides remain.

In principle, any other polymer with at least two different monomers is suitable for information storage as long as efficient encoding and decoding can be achieved (*14*–*16*). Peptides, another major biopolymer, hold great potential as an information storage system (Fig. 1A) (*17*). The enhanced stability of peptides relative to DNA is evidenced by fossil samples in which DNA has already degraded but peptides or proteins can still be detected and sequenced (*18*). Sequencing of peptides can be achieved using nanoscale liquid chromatography–tandem mass spectrometry (nLC-MS/MS), a technique routinely applied in proteomics and de novo peptide discovery (*19*, *20*). Accordingly, nLC-MS/MS can be used to decode information stored in peptides. The encoding of information in peptides can be accomplished through solid-phase peptide synthesis (SPPS), which enables rapid and efficient access to peptides or even proteins (*21*). We envisioned that these technologies may jointly serve as a foundation for the creation of a next-generation information storage system based on peptides as information carriers. Notably, SPPS is compatible with combinatorial synthesis, a key requirement for applications in encoded library platforms. Coupled with the chemical stability of peptides, a peptide-encoded library (PEL) platform would offer greater synthetic versatility than is currently available in encoded libraries (Fig. 1C) (*22*–*25*).

Encoding and decoding information using peptides

Decoding of information stored in a peptide is achieved through the design of an encoding sequence enabling high-fidelity sequencing using an optimized mass spectrometry protocol. We identified a set of 16 nonisobaric amino acids that function as units of information. The canonical and noncanonical amino acids are readily introduced by chemical synthesis with appropriate protection of their side chains (Fig. 2A, encoding monomers). The resulting peptides are sequenced by computational analysis of the secondary mass spectra using high-throughput proteomics data analysis software (*19*). The algorithm matches the fragmented masses observed in secondary mass

[1]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]The Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Center for Environmental Health Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [4]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA.
***Corresponding author. Email: sbuchwal@mit.edu (S.L.B.); blp@mit.edu (B.L.P.)**
†These authors contributed equally to this work.

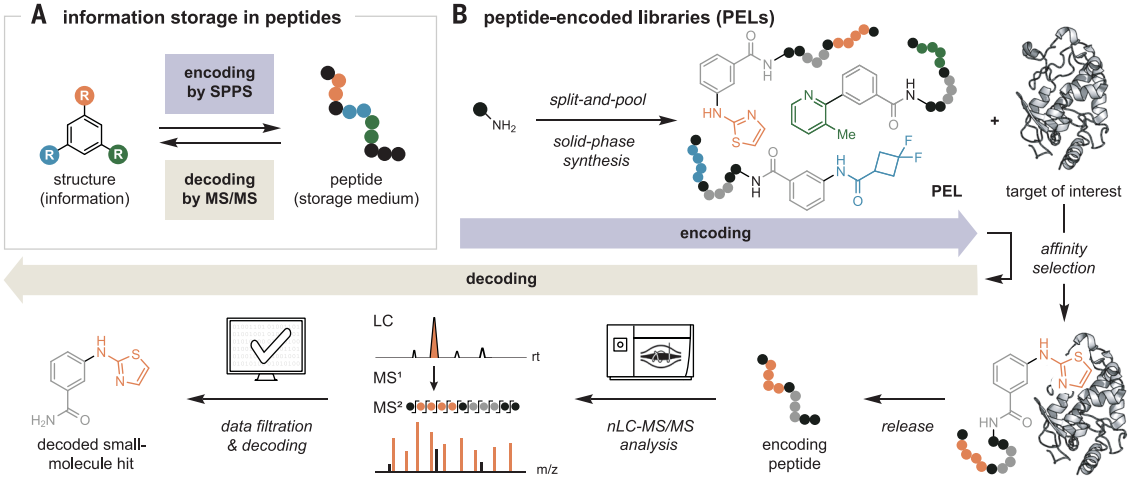


Fig. 1. Peptides are suitable biopolymers for information storage. (**A**) Schematic representation of information storage in peptides. Information such as chemical structures (left) can be encoded in peptides (right) by solid-phase peptide synthesis (SPPS) and decoded using precision sequencing by tandem mass spectrometry (MS/MS). (**B**) Information storage in peptides enables the encoding of synthetic procedures during split-and-pool library synthesis to afford peptide-encoded libraries (PELs), which can be used for affinity-based selections with a target of interest. Decoding of enriched peptide sequences enables the discovery of small-molecule hits for further drug development. m/z, mass-to-charge ratio.

spectra to the predicted fragments of all possible peptides with the defined amino acids in the encoding alphabet. On the basis of the match of the observed ion fragments and the mass error, the software assigns each amino acid in a peptide sequence a local confidence. The sequencing fidelity of the entire peptide is evaluated through the average local confidence (ALC) of each residue of the peptide. Because the sequence and molecular composition of the synthetic peptide are precisely defined, the resulting sets of potential decoded sequences are filtered according to tag sequence design rules to facilitate analysis (*26*). However, prospecting experiments revealed that arbitrary amino acid sequences with variable polarity, charge states, ionization potential, and fragmentation efficiency hampered high-fidelity decoding. Accordingly, the sequencing fidelity of peptide tags was further optimized through the identification of a sequence framework in which the encoding amino acids are embedded. Therein, a peptide tag of 11 amino acids features four spacer monomers that allow fine-tuning of desired tag properties (Fig. 2A). These monomers appear in fixed positions, thus increasing sequencing confidence by functioning as structural restraints during data filtration. Moreover, basic residues enhance the solubility of the resulting tags and increase sequencing accuracy, as represented by higher ALC scores, while protic side chains allow fine-tuning of polarity (*27*). A screening of basic and protic residues at different fixed positions revealed an optimal tag structure featuring a lysine at the C terminus and close to the N terminus, with a serine in the core of the tag and an aliphatic residue on the N terminus [see supplementary materials (SM) section 2.1]. The precisely defined composition of the encoding tag was leveraged for information retrieval during nLC-MS/MS acquisition by targeted mass search (Fig. 2B). Therein, secondary mass spectra are acquired only of ions corresponding to the tag design, which increases the efficiency of analyses by discarding incorrect ions and augments the amount of correctly determined sequences. After targeted mass acquisition, the optimized sequencing protocol relies on data processing wherein sequence filtration by tag design affords an array of plausible sequences. Because a peptide can be selected repeatedly for secondary mass spectrometry during nLC-MS/MS acquisition, a specific tag sequence may occur multiple times in the filtered array. Identical sequences are aggregated and assigned a count based on their number of occurrences as well as an average ALC (aALC) based on the mean of the individual peptide ALC. The optimized tag structure and protocol allow successful sequencing by nLC-MS/MS using as little as 10 fmol of peptide (see SM section 2.2). Using the set of 16 encoding amino acids, millions of pieces of information can be stored in peptide tags with an average molecular weight of <1.5 kDa using an eight-digit encoding string. Theoretically, the high information capacity of a hexadecimal numeral system allows storage of 4.3 billion possible codes in an eight-digit string, whereas binary or quaternary numeral systems enable the storage of only 256 or 56,535 codes, respectively (Fig. 2C).

Information encoded in peptide tags can endure exposure to harsh chemical environments. To probe the stability of information stored in peptide tags as compared with DNA tags, we subjected both types of tags to a procedure reported for the evaluation of the DNA compatibility of chemical reaction conditions (*6*). Accordingly, a peptide tag on polystyrene solid support and a DNA tag on magnetic beads were exposed to various conditions, after which the postexposure integrity of the tags was analyzed by LC-MS and quantitative polymerase chain reaction (qPCR), respectively (Fig. 3; see SM section 2.3). In comparison to a control sample, the peptide tag exhibited superior stability over DNA for all but one of the conditions tested, with >95% of intact peptide tag

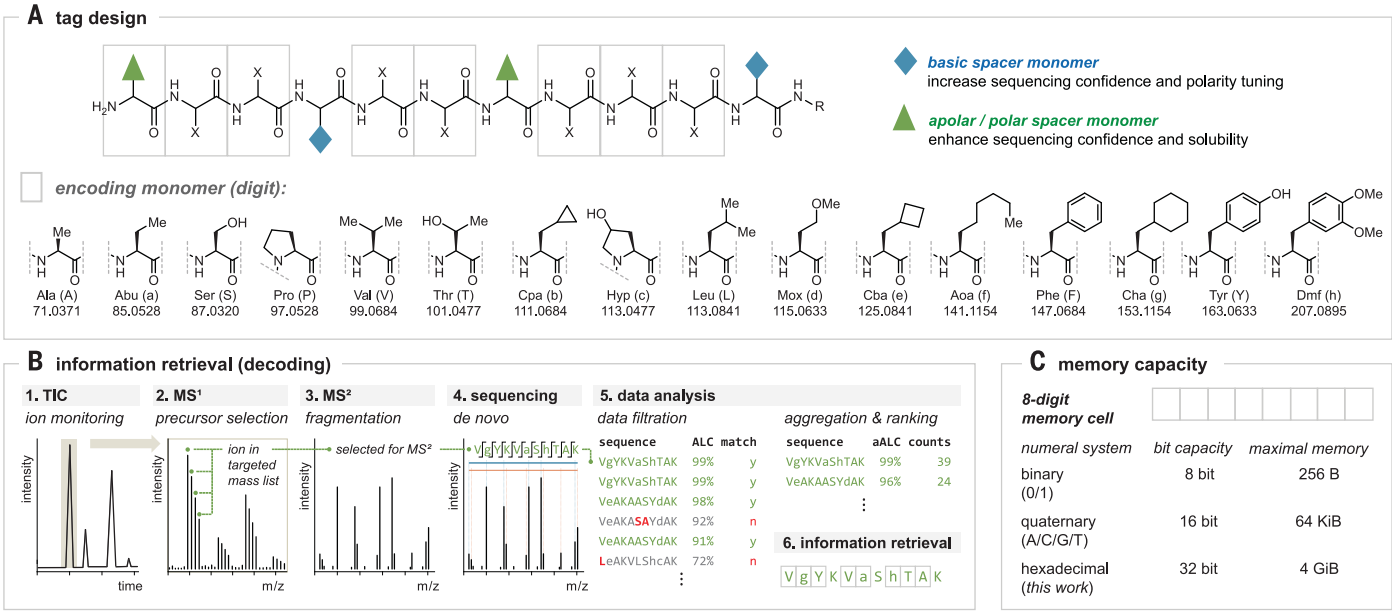
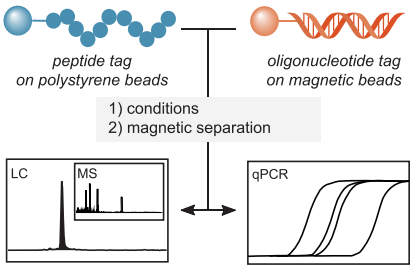


Fig. 2. Peptide tags enable dense memory storage with high information stability. (**A**) Eleven-mer peptide tag composed of coding (X) and spacer (green triangles and blue diamonds) monomers. Encoding positions (gray boxes) are occupied by one of 16 nonisobaric amino acids, labeled with the corresponding three-letter code, one-letter code, and monoisotopic mass (daltons). (**B**) Schematic representation of the decoding workflow. Step 1: The total ion chromatogram (TIC) of the nLC-MS is used to monitor ions of eluting peptide tags by their intensity. Step 2: Primary mass spectra (MS^1) are collected to select ions for fragmentation according to a targeted precursor inclusion list defined by the tag design. Step 3: Secondary mass spectra (MS^2) are generated from fragmentation of selected precursor ions. Step 4: Software-assisted de novo sequencing converts peaks and intensities of MS^2 to peptide sequences on the basis of the specific masses of their encoding monomers. Step 5: The generated sequences are filtered to remove tags not matching the general design and tags with low average local confidence (ALC). Remaining sequences are aggregated and ranked according to their average sequencing confidence (aALC) and the number of sequence counts. Step 6: The tag is spliced according to the defined encoding digits, and the information is retrieved upon decoding the encoding monomers. (**C**) Memory capacity of an eight-digit memory cell with corresponding bit capacity and maximum memory for various numeral systems.



Condition			peptide	DNA
$Pd(OAc)_2$	5%	THF	>95%	<1%
TFA	5%	H_2O	>95%	22%
H_2O_2	5%	H_2O	>95%	>95%
NaOH	0.1 M	H_2O	>95%	22%
$PhI(OAc)_2$	5%	MeCN	>95%	74%
$[Cu(MeCN)_4]PF_6$	5%	MeCN	>95%	88%
Grubbs G2	1%	DCM	>95%	23%
Piperidine	20%	DMF	>95%	17%
Et_3N	5%	DMF	>95%	23%
NaO^{t}Bu	1 M	THF	<1%	16%
UV	254 nm	H_2O	>95%	83%

Fig. 3. Stability of model peptide and DNA tag under synthetically relevant conditions. Suspensions of tags on solid support were subjected to reagents (amounts given as weight percent or molarity) or ultraviolet (UV) irradiation for 4 hours. The peptide and DNA tags were separated and analyzed for integrity by liquid chromatography–mass spectrometry (LC-MS) and quantitative polymerase chain reaction (qPCR), respectively (n = 3 technical replicates). Integrity of the peptide tag is given in area percent of the LC-MS trace, integrity of the DNA tag was calculated from qPCR quantification cycle values as compared with magnetic beads exposed to water only. TFA, trifluoroacetic acid; THF, tetrahydrofuran; DCM, dichloromethane; DMF, *N*,*N*′-dimethylformamide.

detected, whereas most conditions caused severe damage to the DNA tag. The range of conditions suggests that peptide encoding tags can be subjected to broad reaction conditions without compromising the integrity of the stored information, whereas DNA may face limitations under certain chemical conditions. Notably, the amplifiability of DNA is not routinely evaluated in many applications of DELs (*6*).

Synthesis of peptide–small molecule conjugates

We established a molecular scaffold that features two loci for orthogonal synthetic elaboration (Fig. 4A), allowing for sequential small-molecule and peptide synthesis enabled by a fully orthogonal protecting group strategy. The scaffold is bound to a polystyrene bead via a Rink amide linker cleavable under strongly acidic conditions that also result in global deprotection of the conjugate. A lysine residue serves as a branching point to covalently link the peptide and small molecule. The peptide is connected to the scaffold by a Seramox linker orthogonally cleavable under optimized oxidative conditions (see SM section 2.5) (*28*), allowing release of the peptide for sequencing. The peptide and small molecule are functionalized sequentially using orthogonal protection groups (Fmoc, Trt, Alloc), which enables encoding by peptide coupling of specific amino acids either before or after the corresponding small-molecule functionalization. Using this optimized scaffold, we established a synthetic tool kit to allow versatile synthesis of tagged small molecules.

The chemical stability of protected peptides and synthetic utility of solid-phase synthesis allowed the implementation of palladium-mediated cross-coupling reactions (Fig. 4B). Among the top 20 most frequently applied reactions in the medicinal chemistry literature, palladium-catalyzed formation of C–C and C–N bonds represent the only reactions discovered in the past four decades (*29*). The possibility of generating molecular diversity through the coupling of two building blocks renders these reactions particularly suited for the exploration of novel chemical space and attractive for small-molecule library synthesis. Peptide-based encoding tags allow the use of reaction conditions optimized to promote high yields and selectivity, rather than reaction conditions contrived to meet the stability demands of DNA. For the palladium-mediated formation of C–C bonds, we found that a fourth-generation palladium precatalyst featuring biaryl phosphine ligand XPhos (XPhos Pd G4) (*30*) enables the cross-coupling of widely available aryl boronic acids with resin-bound aryl bromide–peptide conjugates. Under basic conditions using mixed organic aqueous reaction media at room temperature, 36 boronic acids and 14 aryl bromides were found to undergo cross-coupling to afford the desired product with purity deemed sufficient for inclusion in library synthesis (>70% purity; see SM section 2.7 for scope). The cross-coupling of aniline derivatives to resin-bound bromide–peptide conjugates to forge C–N bonds was achieved using an AlPhos-ligated palladium dimer and DBU in THF at 50°C (*31*). Under these optimized conditions, 41 anilines and 11 aryl bromides were found to undergo cross-coupling with >70% purity and were selected for inclusion in library synthesis (see SM section 2.8 for scope). In both cross-coupling reactions, any remaining material mostly corresponds to product resulting from protodemetalation. The scope and purity of these reactions outperform recent examples of palladium-mediated cross-couplings in the presence of DNA (*32*–*34*), allowing for the cross-coupling of diverse heterocycles prevalent in drugs (selected examples shown in Fig. 4C; for full scope see SM section 2.8). Additionally, 60 carboxylic acids and acid chlorides were found to be competent substrates for amide coupling, affording the corresponding products with excellent purity (see SM section 2.9). The compatibility of peptide tags with different synthetic conditions was further established through implementation of an acid-mediated Pictet-Spengler reaction (see SM section 2.10 for details), which has been reported to result in loss of encoding information even in stabilized DEL systems (*9*). These reactions together with our optimized protecting group strategies relying on palladium or trifluoroacetic acid for deprotection establish the broad compatibility of peptide tags with a range of synthetic conditions. Moreover, products of each reaction were confirmed to efficiently undergo subsequent elaboration of the peptide tag, thus enabling the encoding of chemical information.

Peptide-encoded small-molecule libraries

Two peptide-encoded libraries (PELs) of small molecules were prepared using combinatorial chemistry. We designed a library synthesis based on a split-and-pool strategy allowing orthogonal encoding and small-molecule elaboration using our optimized reactions with building blocks confirmed to yield high purity (see SM 1.5 for detailed synthesis scheme). Accordingly, during the split step of combinatorial PEL synthesis, a building block is coupled to the small molecule and encoded through either previous or subsequent coupling of corresponding amino acids. The libraries featured central building blocks (BB1) bearing a protected amine for coupling of carboxylic acid building blocks (BB2) as well as an aryl bromide for palladium-mediated cross-coupling reactions with amines or boronic acids (BB3), affording a 41K-membered C–N or 39K-membered C–C-based library, respectively. Each building block was encoded during synthesis using a randomly assigned combination of amino acids, resulting in a unique, defined peptide sequence for each library member. Encoding each building block allowed the incorporation of regio- and stereoisomers, thus achieving chemical diversity not achievable in standard mass-based discovery platforms (*35*). Because libraries are mixtures of compounds and thus not suitable for purification or purity assessments, high-yielding reaction sequences are imperative to robust downstream applications. To confirm the viability of the synthetic sequence in our library design, we first synthesized individual library members under the synthetic conditions applicable to library synthesis (Fig. 4D). We found that completing 45 synthetic manipulations according to the library design, including global deprotection and resin cleavage, afforded the desired products resulting from C–C and C–N cross-coupling in crude purities of >70%. Notably, the efficiency of solid-phase synthesis enables such multistep library synthesis to be carried out in less than a week. Quantitative estimation of drug likeness (*36*)

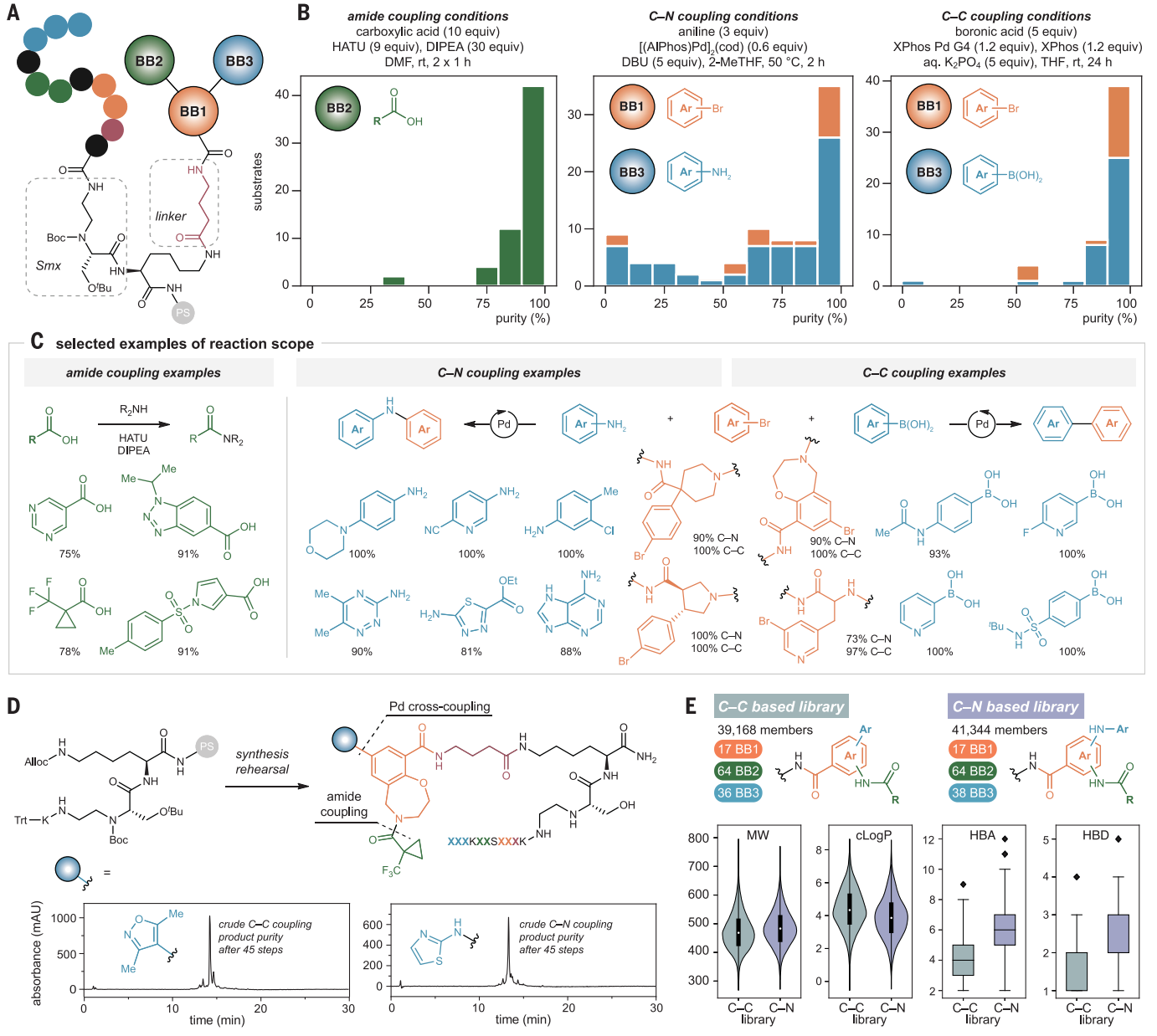


Fig. 4. Small molecule–peptide conjugates formed from three building blocks and 11 encoding amino acids provide a suitable scaffold for the implementation of diversity-oriented reactions. (**A**) Schematic representation of the library design. The small molecule formed from three building blocks (BB1, BB2, and BB3) is linked to the encoding peptide through a short aliphatic linker, a branching lysine tethered to solid support (PS, polystyrene), and an oxidatively cleavable linker (Smx). (**B**) Reaction outcome for amide coupling, and palladium-mediated C–C and C–N bond-forming reaction. The corresponding histograms show the distribution of purities obtained from coupling of different building blocks determined by LC-MS integration. The amide coupling histogram includes four acid chlorides that were coupled using DIPEA in DCM. HATU, 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate; DIPEA, *N*,*N*-diisopropylethylamine; DMF, *N*,*N*′-dimethylformamide; DBU, 1,8-diazabicyclo(5.4.0)undec-7-ene; 2-MeTHF, 2-methyltetrahydrofuran; AlPhos, 6-methoxy-2-(2,4,6-tri-i-propyl-3-((2,3,5,6-tetrafluoro-4-butyl)phenyl)phenyl)diadamantyl phosphine; XPhos, 2-dicyclohexylphosphino-2′,4′,6′-triisopropylbiphenyl; rt, room temperature. (**C**) Selected examples demonstrate the scope of the optimized coupling reactions. Percent purity, as determined by LC-MS integration. (**D**) Crude purity of model compounds after 45-step synthesis rehearsal. (**E**) Calculated molecular properties of the small molecules contained in PELs resulting from C–C or C–N cross-coupling reactions. Molecules were modeled with the linker terminating as a methyl amide to afford distributions of molecular weight (MW), lipophilicity as the calculated partition coefficient (cLogP), and number of hydrogen bond acceptors (HBA) and donors (HBD). MW and cLogP distribution are shown as a kernel density estimation with a median (white dot), interquartile range (bold black line), and 1.5× interquartile range (black line). HBA and HBD are illustrated with an interquartile range (box), median (horizontal line), full range (black bar), and outliers (black diamonds).

of the individual library members showed that most compounds exhibit molecular properties advantageous for lead drug candidates (Fig. 4E) (*37*). Random sampling and sequencing of peptide libraries corresponding to the encoding peptides as found in the PELs revealed low sequencing bias of individual amino acids in this PEL design (see SM section 2.13.2).

The affinity selection workflow with PELs was optimized using carbonic anhydrase IX (CA IX). CA IX is a metalloenzyme involved in tumor acidosis and a promising oncology

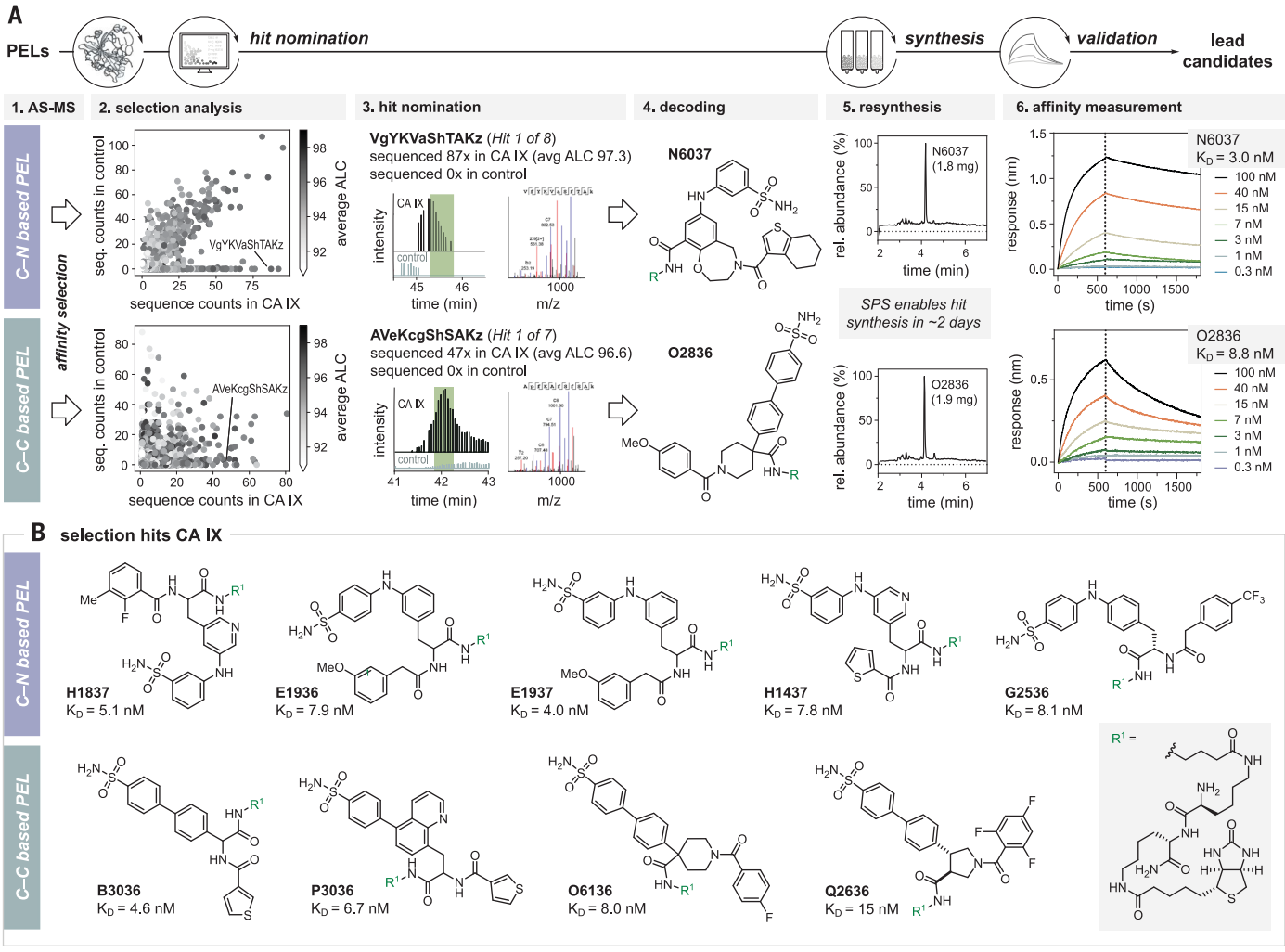


Fig. 5. PELs enable the discovery of novel small molecules with high affinity for CA IX. (**A**) Schematic representation of PEL discovery workflow. Step 1: Affinity selection–mass spectrometry (AS-MS). Step 2: Selection analysis showing aggregated sequence counts in selections against target protein (CA IX; n = 3 replicates) and control (streptavidin; n = 3 replicates). Gray sequential colormap illustrates average of ALC of each sequence count in the target protein. Step 3: Hit nomination. Extracted ion chromatograms (EICs) of MS^1 and MS^2 for the highlighted sequences are shown. Step 4: Decoding. The peptide sequence is decoded to give the corresponding small molecule (R, linker). Step 5: Resynthesis. Solid-phase synthesis (SPS) enables access to milligram quantities of potential hits within days. Step 6: Affinity measurement. Dissociation constant (K_D) is determined by biolayer interferometry (BLI). (**B**) Molecular structures of nominated hits from the C–N–based (top row) and C–C–based (bottom row) PELs (R^1, linker-biotin). K_D values were determined by BLI.

target (*38*), which has previously been reported as a useful protein for DEL development owing to favorable binding interactions of sulfonamides with a zinc-featuring active site (*39*). Accordingly, optimization of the affinity selection conditions and workflow was carried out using a 400-membered PEL featuring sulfonamide-bearing small molecules corresponding to the PEL design in Fig. 4A. Evaluation of incubation concentration, conditions for cleavage of the peptide tag, and desalting procedures resulted in optimized affinity selection conditions that afforded consistent enrichment of peptide sequences encoding sulfonamide-featuring library members. A peptide sequence resulting from a conjugate with a strong enrichment was selected for further evaluation. The corresponding small molecule as well as the small molecule–encoding peptide conjugate were confirmed to exhibit nanomolar affinity for CA IX. Conversely, the encoding peptide displayed no affinity by itself (see SM section 2.14).

The 41K-membered C–N and 39K-membered C–C-based PELs were used in affinity selections for the discovery of small molecules with nanomolar affinity for CA IX. Using the optimized affinity selection protocol, both PELs were screened against immobilized CA IX (Fig. 5A, step 1). In an automated selection procedure, biotinylated CA IX was immobilized on streptavidin-featuring magnetic beads and incubated with a PEL. Subsequently, the removal of nonbinding members was achieved through repeated washing steps. Encoding peptides were released from retained conjugates under oxidative conditions and subjected to analysis by nLC-MS/MS with targeted mass acquisition. This procedure was conducted in replicates for the target protein as well as unfunctionalized streptavidin beads as a control for binding specificity. After sequencing of the encoding peptide tags, the data for all replicates was filtered according to the library design and aggregated to enable ranking of encoding peptides by their sequence counts in the target and control protein (step 2). Sequencing confidence, as represented by the corresponding average ALC and target specificity, were taken into consideration for subsequent hit nomination (step 3). Decoding of the

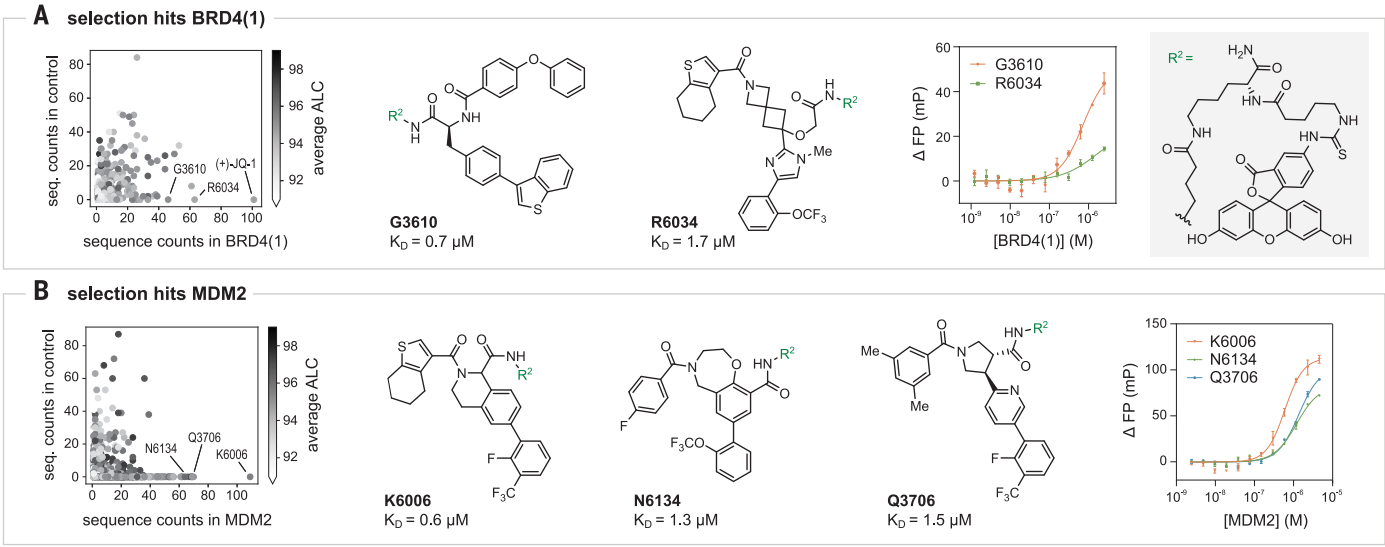


Fig. 6. PELs enable the discovery of novel small molecules with high affinity for their respective target of interest. (**A**) Affinity selection results from C–C–based PEL screen against BRD4(1). Enrichment plot (left), nominated hits (R^2, linker-fluorescein as depicted in the gray box) with K_D determined by fluorescence polarization (FP), and concentration-dependent FP (right). (**B**) Affinity selection results from C–C–based PEL screen against MDM2. Enrichment plot (left), nominated hits [R^2, linker-fluorescein as depicted in (A)] with K_D determined by FP, and concentration-dependent FP (right).

peptide tags to their corresponding small molecules afforded several candidates (step 4), which were further filtered for their calculated drug-like properties. Eleven hit molecules were selected and synthesized expediently on solid support on a time scale of 2 days (step 5), requiring only a single purification. The individually synthesized small molecules were subjected to validation by biolayer interferometry (BLI, step 6). All hit molecules were found to exhibit affinities for CA IX in the single-digit or low double-digit nanomolar range (Fig. 5B), which renders them promising lead candidates for further evaluation. Additionally, peptide libraries without appendant small molecules corresponding to the encoding peptides of the C–N and C–C–based PELs were synthesized and subjected to affinity selection with CA IX (see SM section 2.15.2). No enrichment was observed, suggesting negligible contribution to binding by the encoding peptides.

The applicability of PELs in hit discovery was further established through de novo discovery of small molecules with nano- to micromolar affinity for proteins involved in oncogenic pathways. Bromodomain-containing protein 4 (BRD4) represents a transcriptional regulator involved in tumor growth with a reported ligand, (+)-JQ-1 (*40*). We subjected the C–N and C–C–based PELs to affinity selection against immobilized BRD4(1), the bromodomain 1 of BRD4. To test the enrichment of a high-affinity ligand, a (+)-JQ-1 conjugate featuring a corresponding encoding peptide was included in the affinity selection. After sequencing and data filtration, the peptide encoding (+)-JQ-1 was found to exhibit a high sequence count for BRD4(1) (>100 counts) (Fig. 6A). Additionally, two peptide tags from the C–C–based PEL were found to exhibit >40 sequence counts with >94% average ALC. Synthesis of the corresponding small molecules and validation by fluorescence polarization (FP) confirmed their affinity to BRD4(1). We additionally subjected the PELs to affinity selection against MDM2, an E3 ubiquitin ligase involved in tumor proliferation (*41*). Seven hits from the C–C–based PEL were selected on the basis of sequence counts and ALC, synthesized, and confirmed to exhibit affinity for MDM2 using FP (Fig. 6B and SM section 2.15.6.).

Discussion

We herein describe the use of abiotic peptides as an information storage system with robust encoding through synthesis and decoding by nLC-MS/MS peptide sequencing. This encoding strategy bears distinct advantages. Amino acids selected for the encoding alphabet must meet certain requirements such as chemical stability, low sequencing bias, and defined charge states. Because SPPS for encoding and mass spectrometry–based decoding allows for the inclusion of noncanonical amino acids, an encoding alphabet can be curated from a vast pool of possible nonisobaric amino acids, including isotopologues. We established a hexadecimal numeral system, which results in a theoretical information density exceeding that of binary and quaternary systems, for example, DNA. Although the use of noncoding monomers to further enhance sequencing fidelity leads to a decrease in the achievable net information capacity, similar effects occur in DNA as a result of primers and barcoding demands (*14*).

The encoding strategy relying on abiotic peptides additionally affords several advantages from the perspective of synthetic chemistry. A core strength results from the ability to use reaction conditions chosen for their efficiency instead of by the requirements of the encoding tag. Moreover, solid-phase synthesis offers the advantage of allowing the use of a large excess of a reagent to drive reaction efficiency and avoiding laborious workups and purification procedures. Accordingly, we established high-yielding reactions for the synthesis of amides, anilines, and biphenyls, which constitute functional groups appearing in ~50, 40, and 10% of entries, respectively, in a database of 6.2 million pharmaceutical drug discovery patents and journals (*29*). Optimized reaction conditions for small-molecule transformations enable the inclusion of demanding substrates, as exemplified by the palladium-mediated cross-coupling reactions featuring complex heteroaryl substrates. Together, the reaction efficiency results in high-purity peptide–small molecule conjugates, a key requirement in subsequent information applications. The high purity of the conjugates stands in stark contrast to DELs, where decoded hits often require triaging through resynthesis under conditions used for library preparation to identify the active compound among numerous unintended products (*42*). However, even in the event of the formation of an unexpected product during PEL synthesis, hit resynthesis is conducted using reaction conditions analogous to those used

during library synthesis and would consequently afford the corresponding hit.

We demonstrated the successful de novo discovery of small-molecule ligands for CA IX, BRD4(1), and MDM2 from PELs through affinity selections. The zinc-bearing active site of CA IX results in favorable interactions with sulfonamides, characterized by fast k_{on} and slow k_{off} rates (*43*), which render the protein a popular target in platform development (*39*). Whereas the evaluation of proteins with known binding motifs can hedge against the uncertainty associated with library discovery, the application of PELs against additional protein targets establishes broader applicability of PELs for the identification of small-molecule protein binders, demonstrated herein for BRD4(1) and MDM2. Promising future avenues include discovery campaigns focused on transcription factors or nonprotein targets such as RNA. Although DELs have successfully been screened against RNA structures (*44*), combinatorial libraries with abiotic peptide tags are expected to show lower nonspecific interaction with oligonucleotides and could therefore be a valuable, orthogonal discovery platform for such emerging classes of targets.

The peptide-based encoding tag in PELs can be sequenced reliably at quantities suitable for affinity selections. However, the limit of detection is the subject of ongoing optimization, as it is expected to correlate with the utilizable library size. The observed sequencing threshold of as little as 10 fmol of peptide suggests that larger libraries are feasible, but we have not explored the upper limit of PEL size in this study. Unlike DNA encoding tags, which can be amplified by PCR before sequencing, peptide tags cannot be amplified by equivalent biochemical methods. Although the tag amount used for high-fidelity decoding in this study exceeds amounts reported previously for DELs (*45*), we anticipate that continuous improvement of mass spectrometric instrumentation and methods will substantially reduce the sample amount required for reliable decoding of peptide tags in the future.

In summary, the findings described herein demonstrate the use of peptides as an information storage medium for small-molecule synthesis. These small molecule–peptide conjugates enable the use of diverse transformations for small-molecule synthesis and can be translated toward a new drug discovery platform called a peptide-encoded library. We expect this technology to have implications for biochemical data storage technology, peptide–small molecule synthesis, and drug discovery.

Methods summary

Detailed information on instrumentation, general procedures, supplemental results, and analytical data can be found in the supplementary materials.

Synthesis

PELs were prepared by combinatorial synthesis on solid support (polystyrene resin) employing reactions and building blocks previously validated to yield high-purity products using a model synthesis scaffold (detailed synthetic scheme SM 1.5). During combinatorial split-and-pool synthesis, the resin is split into a number of wells corresponding to the number of building blocks for the specific step. In each well, the corresponding building block is coupled to the small-molecule fragment and encoded through previous or subsequent coupling of the corresponding amino acids. These sequential reactions are enabled through the use of orthogonal protecting groups. Upon completed functionalization and encoding, the resin is pooled. Pd-mediated C–C cross-coupling was conducted using aryl bromide on resin, the corresponding boronic acid (5 equiv), Pd G4 XPhos (1.2 equiv) and XPhos (1.2 equiv) in THF, and aqueous K_3PO_4 (0.5 M, 5 equiv) for 24 hours at room temperature. Pd-mediated C–N cross-coupling was conducted under N_2 atmosphere using aryl bromide on resin, aniline (3 equiv), DBU (5 equiv), and [$(AlPhos)_2Pd_2(cod)$] (0.6 equiv) in 2-MeTHF at 50°C for 2 hours. The amide coupling for small-molecule synthesis was conducted by treatment of amine on resin with carboxylic acid (10 equiv), HATU (9 equiv), and DIPEA (30 equiv) in DMF twice for 1 hour.

Affinity selection

Biotinylated proteins were immobilized on streptavidin-coated magnetic beads. After incubation of the protein of interest with the PELs, the protein-binder complex was washed to remove nonbinding members of the library. The encoding peptide tags were released by oxidative cleavage using sodium periodate, concentrated, and desalted for analysis by nLC-MS/MS. The procedure was performed in parallel using streptavidin-coated magnetic beads with no added protein to control for binding specificity.

nLC-MS/MS

The concentrated, desalted samples of the protein of interest and the streptavidin control were subjected to analysis by nLC-MS/MS on an Orbitrap Eclipse Tribrid mass spectrometer. Two secondary mass spectra were acquired upon fragmentation by higher-energy collisional dissociation (HCD) and electron-transfer dissociation with supplemental activation by HCD (EThcD) of ions from a targeted mass inclusion list corresponding to the expected ions of the PEL tags.

Hit nomination

Raw data from nLC-MS/MS were sequenced by PEAKS (*46*), a proteomics software for de novo peptide sequencing of secondary mass spectra, by defining the noncanonical amino acids as posttranslational modifications. The sequences identified therein were filtered to retain tags corresponding to the library design. Subsequently, the tags were ranked by their aggregated sequence counts in the selection replicates of the target of interest and the control samples. Small molecules corresponding to peptide sequences with high counts for the target protein and low counts in the control were selected for further evaluation. Additional considerations can include calculated molecular properties or manual evaluation of MS^1 and MS^2 spectra.

Hit synthesis and validation

Small molecules selected for hit validation were synthesized on solid support using the reaction conditions described above. The small molecules were synthesized with an appendant biotin for BLI or fluorescein for FP. Affinity for a target protein was validated by variable concentration assays using BLI or FP.

REFERENCES AND NOTES

1. J. Shendure *et al*., *Nature* **550**, 345–353 (2017).
2. L. Ceze, J. Nivala, K. Strauss, *Nat. Rev. Genet.* **20**, 456–466 (2019).
3. D. Neri, R. A. Lerner, *Annu. Rev. Biochem.* **87**, 479–502 (2018).
4. R. E. Kleiner, C. E. Dumelin, D. R. Liu, *Chem. Soc. Rev.* **40**, 5707–5717 (2011).
5. R. A. Goodnow Jr., C. E. Dumelin, A. D. Keefe, *Nat. Rev. Drug Discov.* **16**, 131–147 (2017).
6. M. L. Malone, B. M. Paegel, *ACS Comb. Sci.* **18**, 182–187 (2016).
7. J. H. Hunter *et al*., *J. Org. Chem.* **86**, 17930–17935 (2021).
8. D. T. Flood *et al*., *ACS Cent. Sci.* **6**, 1789–1799 (2020).
9. M. Potowski *et al*., *Angew. Chem. Int. Ed.* **60**, 19744–19749 (2021).
10. A. B. MacConnell, P. J. McEnaney, V. J. Cavett, B. M. Paegel, *ACS Comb. Sci.* **17**, 518–534 (2015).
11. C. Zambaldo, S. Barluenga, N. Winssinger, *Curr. Opin. Chem. Biol.* **26**, 8–15 (2015).
12. P. Blakskjaer, T. Heitner, N. J. V. Hansen, *Curr. Opin. Chem. Biol.* **26**, 62–71 (2015).
13. Y. Ruff *et al*., *ACS Comb. Sci.* **22**, 120–128 (2020).
14. M. G. T. A. Rutten, F. W. Vaandrager, J. A. A. W. Elemans, R. J. M. Nolte, *Nat. Rev. Chem.* **2**, 365–381 (2018).
15. B. J. Cafferty *et al*., *ACS Cent. Sci.* **5**, 911–916 (2019).
16. C. Cao *et al*., *Sci. Adv.* **6**, eabc2661 (2020).
17. C. C. A. Ng *et al*., *Nat. Commun.* **12**, 4242 (2021).
18. M. Warren, *Nature* **570**, 433–436 (2019).
19. H. Steen, M. Mann, *Nat. Rev. Mol. Cell Biol.* **5**, 699–711 (2004).
20. A. J. Quartararo *et al*., *Nat. Commun.* **11**, 3183 (2020).
21. N. Hartrampf *et al*., *Science* **368**, 980–987 (2020).
22. J. M. Kerr, S. C. Banville, R. N. Zuckermann, *J. Am. Chem. Soc.* **115**, 2529–2531 (1993).
23. J. Vágner *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8194–8199 (1996).
24. R. Liu, J. Marik, K. S. Lam, *J. Am. Chem. Soc.* **124**, 7678–7680 (2002).
25. A. Song, J. Zhang, C. B. Lebrilla, K. S. Lam, *J. Am. Chem. Soc.* **125**, 6180–6188 (2003).
26. A. A. Vinogradov *et al*., *ACS Comb. Sci.* **19**, 694–701 (2017).
27. V. H. Wysocki, G. Tsaprailis, L. L. Smith, L. A. Breci, *J. Mass Spectrom.* **35**, 1399–1406 (2000).
28. S. Pomplun *et al*., *Angew. Chem. Int. Ed.* **59**, 11566–11572 (2020).
29. D. G. Brown, J. Boström, *J. Med. Chem.* **59**, 4443–4458 (2016).
30. N. C. Bruno, N. Niljianskul, S. L. Buchwald, *J. Org. Chem.* **79**, 4161–4166 (2014).
31. J. M. Dennis, N. A. White, R. Y. Liu, S. L. Buchwald, *J. Am. Chem. Soc.* **140**, 4721–4725 (2018).

32. Y.-C. Chen *et al*., *Bioconjug. Chem.* **31**, 770–780 (2020).
33. X. Lu, S. E. Roberts, G. J. Franklin, C. P. Davie, *MedChemComm* **8**, 1614–1617 (2017).
34. P. R. Chheda, N. Simmons, D. P. Schuman, Z. Shi, *Org. Lett.* **24**, 3401–3406 (2022).
35. R. Prudent, D. A. Annis, P. J. Dandliker, J.-Y. Ortholand, D. Roche, *Nat. Rev. Chem.* **5**, 62–71 (2021).
36. G. R. Bickerton, G. V. Paolini, J. Besnard, S. Muresan, A. L. Hopkins, *Nat. Chem.* **4**, 90–98 (2012).
37. C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Deliv. Rev.* **46**, 3–26 (2001).
38. S. Pastorekova, R. J. Gillies, *Cancer Metastasis Rev.* **38**, 65–77 (2019).
39. F. Buller *et al*., *ACS Chem. Biol.* **6**, 336–344 (2011).
40. P. Filippakopoulos *et al*., *Nature* **468**, 1067–1073 (2010).
41. M. Konopleva *et al*., *Leukemia* **34**, 2858–2874 (2020).
42. W. Su *et al*., *Bioconjug. Chem.* **32**, 1001–1007 (2021).
43. C. T. Supuran, *Nat. Rev. Drug Discov.* **7**, 168–181 (2008).
44. R. I. Benhamou *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2114971119 (2022).
45. A. Sannino *et al*., *ChemBioChem* **20**, 955–962 (2019).
46. B. Ma *et al*., *Rapid Commun. Mass Spectrom.* **17**, 2337–2342 (2003).

ACKNOWLEDGMENTS

We thank A. Callahan, J. Margarini, G. Porter, C.-P. Hsu, A. Loas, and C. Nguyen for their assistance and discussions. E. Miller, M. Strauss, M. Richter, R. Sarott, and B. Schreib are gratefully acknowledged for advice and assistance in the preparation of this manuscript. **Funding:** S.L.R. and N.M.G. acknowledge support from the Swiss National Science Foundation (Early Postdoctoral Mobility Fellowships PN 195700 and PN 195351, respectively). Partial support of research reported in this publication was provided by the National Institutes of Health under award R35GM122483 (S.L.B.). The Orbitrap Eclipse Tribrid mass spectrometer used in this study was acquired with generous funding provided by the Massachusetts Life Sciences Center through the 2020 Novel Therapeutics Delivery program. We thank Millipore-Sigma for the generous gift of XPhos. **Author contributions:** All authors conceptualized the research and wrote the manuscript. S.L.R. and N.M.G. designed, performed, and evaluated the experiments. **Competing interests:** The authors declare the following competing financial interests: MIT filed a patent application (PCT/US22/51802) on the technology described in this study, for which S.L.R., N.M.G., S.L.B., and B.L.P. are listed as inventors. MIT has obtained patents on some ligands and precatalysts described in this manuscript, from which S.L.B. and former co-workers receive royalty payments. **Data and materials availability:** All data generated during this study are available either in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf1354
Materials and Methods
Supplementary Results
Figs. S1 to S43
Tables S1 to S26
References (*47–49*)
MDAR Reproducibility Checklist

Submitted 30 September 2022; accepted 24 January 2023
10.1126/science.adf1354

CONSERVATION ECOLOGY

Ecosystem-based management outperforms species-focused stocking for enhancing fish populations

Johannes Radinger[1]*†, Sven Matern[1,2]†, Thomas Klefoth[3], Christian Wolter[1], Fritz Feldhege[1,2], Christopher T. Monk[1,4], Robert Arlinghaus[1,2,5]

Ecosystem-based management is costly. Therefore, without rigorously showing that it can outperform traditional species-focused alternatives, its broad-scale adoption in conservation is unlikely. We present a large-scale replicated and controlled set of whole-lake experiments in fish conservation (20 lakes monitored over 6 years with more than 150,000 fish sampled) to examine the outcomes of ecosystem-based habitat enhancement (coarse woody habitat addition and shallow littoral zone creation) versus a widespread, species-focused alternative that has long dominated fisheries management practice (i.e., fish stocking). Adding coarse woody habitats alone did not, on average, enhance fish abundance, but creating shallow water habitat consistently did, especially for juvenile fish. Species-focused fish stocking completely failed. We provide strong evidence questioning the performance of species-focused conservation actions in aquatic ecosystems and instead recommend ecosystem-based management of key habitats.

There is a long-standing debate on how effectively ecosystem-based management can counter biodiversity loss, aid in the conservation of imperiled species (*1*), or sustain and rebuild fisheries (*2*, *3*). Ecosystem-based management targets improving or reinstalling key ecological processes, habitats, and species interactions rather than focusing on removing single stressors or supporting individual species (*1*). Globally, the application of ecosystem-based management is still in its infancy (*2*, *3*), and constraints include the strong political and financial support needed (*2*, *4*). Garnering such support is challenging when there are still many unknowns about its effectiveness. Habitat management actions can fail, particularly if they are not wide-ranging enough or do not address key bottlenecks critical in the life cycle of an organism (*5*). Further, experimenting at the scale of natural ecosystems in a replicated fashion is rarely done because it is often practically infeasible or too costly (*6*). Ecosystem-based habitat management may be systematically more effective and sustainable than traditional single-species–oriented measures for achieving conservation objectives because of its comprehensive consideration of the interconnections among species, their environment, and humans (*1*). However, policy-makers are unlikely to fully support implementing such practices on a large scale until robust supporting evidence accumulates.

When species decline, a common species-focused mitigation measure receiving substantial stakeholder and political support is releasing wild-captured or hatchery-bred animals (*7*–*9*); in fisheries, this practice is known as stocking (*8*, *9*). However, stocking can have lasting negative ecological and evolutionary effects on populations, food webs, and ecosystems, e.g., due to the spread of non-native genotypes or species (*8*–*11*). Further, models and empirical studies have demonstrated that releasing fish often fails to increase populations (*8*, *12*–*14*). Nevertheless, species-focused management through stocking in inland fisheries and fish conservation continues to be a standard practice because of a range of psychological (e.g., norms and habits) and institutional (e.g., lack of monitoring) factors (*9*, *15*).

Ecosystem-based approaches to habitat management are promising alternatives to stocking (*9*, *16*). To be successful and to support wild-living animals, management interventions must effectively remediate current population constraints. For fish, key bottlenecks in population dynamics relate to density-dependent mortality in early juveniles, a critical life stage that determines year-class strength and adult abundance (Fig. 1) (*13*, *14*). In particular, the smallest length classes of fish face an important trade-off between securing sufficient food resources to support growth beyond their predators' gape width and minimizing exposure to predation (*17*). This trade-off is shaped by intra- and interspecific competition and the arrangement of profitable foraging areas with higher predation risk versus refuges (e.g., vegetation, structural habitats, or shallow water) that offer protection at the expense of food intake (*18*, *19*) (Fig. 1). Thus, habitat enhancement can improve the growth-mortality trade-off to allow a greater number of juvenile fish to grow into the population, whereas stocking fish into naturally reproducing populations can increase competition or predation without providing refuges from mortality (*14*).

Shallow littoral zone creation (e.g., by excavating new shallow areas) is an ecosystem-based management action that holds promise for remediating habitat constraints (*20*) and effectively addressing the growth-mortality trade-off in juvenile fish. For many fish species, shallow lake zones provide valuable spawning and nursery habitats (*21*) and constitute foraging areas that spatially overlap with safe refuge areas within submerged macrophytes, thereby contributing to fish recruitment and productivity (*22*). An alternative strategy is directly managing habitat structure by introducing coarse woody habitats (*23*, *24*), an important functional habitat for different life stages in many fish species (*25*). However, in lakes, it has remained unclear whether adding coarse woody habitats can effectively increase fish abundance through either improved reproduction (*26*) or provision of refuge benefits (*23*, *27*) that reduce juvenile mortality (*28*), or if the practice simply alters fish distributions by attraction effects (*29*) and habitat partitioning (*23*) without increasing overall abundance (*23*, *30*).

Previous studies have addressed selected aspects of the ecology and conservation value of fish stocking and habitat enhancement in lakes [e.g., (*12, 23*)]. However, lack of controls and insufficient replication (*6*) [but see (*31*)] have limited inference regarding the success of these management measures on broader scales (*32*). Whole-lake experiments (*6*, *23*) have a large potential to systematically evaluate ecosystem-based habitat enhancements versus species-focused stocking, particularly when conducted in a before-after-control-impact (BACI) design (*33*). Small freshwater ecosystems offer excellent opportunities for experimentation and replication (*32*).

We present a large-scale replicated and controlled set of whole-lake experiments in a transdisciplinary setting with strong participatory involvement of local angling communities (*34*). Using 20 mesotrophic gravel pit lakes (average size 7 ha; table S1), we tested for the potential for fish abundance–enhancing effects of three types of management measures: fish stocking with five species in four lakes, habitat enhancement through additions of coarse wood bundles in eight lakes, and shallow littoral zone creation by excavation of riparian banks in a subset of four wood-supplemented lakes (*34*) (figs. S1 to S3). We used a BACI experimental design (including eight control lakes) and monitored the fish community over 6 years to test the following hypotheses: (i) that creating shallow littoral habitats would most effectively

[1]Department of Fish Biology, Fisheries and Aquaculture, Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany. [2]Division of Integrative Fisheries Management, Faculty of Life Sciences, Humboldt-Universität zu Berlin, Berlin, Germany. [3]Ecology and Conservation, Faculty of Nature and Engineering, Hochschule Bremen, Bremen, Germany. [4]GEOMAR Helmholtz Centre for Ocean Research Kiel, Marine Evolutionary Ecology, Kiel, Germany. [5]Integrative Research Institute on Transformations of Human-Environmental Systems (IRI THESys), Humboldt-Universität zu Berlin, Berlin, Germany.
*Corresponding author. Email: johannes.radinger@igb-berlin.de
†These authors contributed equally to this work.

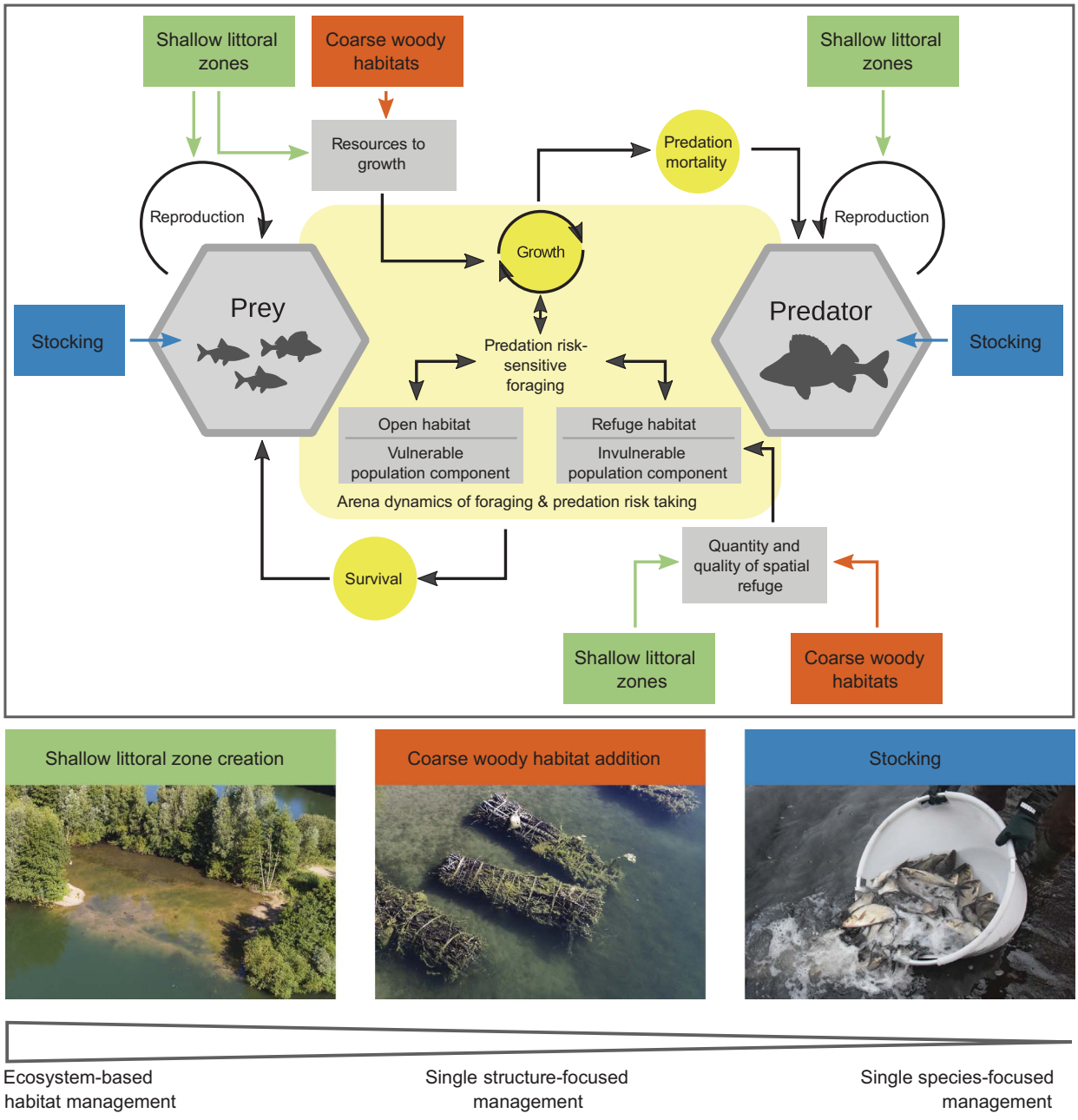


Fig. 1. Population dynamic mechanisms emerging from ecosystem-based habitat management through creating shallow littoral zones and coarse woody habitat addition and from species-focused stocking management. In fish, population constraints related to the density-dependent mortality bottleneck in early juveniles are especially important (*13*, *14*). The yellow box indicates the foraging arena involved in the trade-off between fish growth and mortality in littoral zones of lakes. The smallest length classes of fish face an important trade-off between securing sufficient food resources to support growth to lengths beyond their predators' gape width while minimizing exposure to predation (*17*). Therefore, they adjust their foraging behavior based on the perceived risk of predation (i.e., predation risk–sensitive foraging). This trade-off between predation risk taking and foraging is shaped by intra- and interspecific competition and the spatial distribution of food resources. Juvenile fish may forage in profitable but risky areas outside the refuge (i.e., the vulnerable population component) or move into refuges (e.g., vegetation, coarse woody structures, or shallow water) to limit mortality (i.e., the invulnerable population component) but at the expense of food intake (*18*, *19*). Although habitat enhancement can improve this growth-mortality trade-off to allow a greater number of juvenile fish growing into the population, stocking of otherwise naturally reproducing species would not modify the spatial configuration of foraging arenas; it would only elevate competition or predation without an opportunity to find refuge from mortality. Left and center photos courtesy of Florian Möllers/AVN.

increase fish abundance by providing additional spawning and nursery grounds while simultaneously reducing predation risk because of the beneficial spatial interspersion of foraging with refuge habitats; (ii) that adding coarse woody habitats would create fish aggregations, and a simultaneous attraction of predators and prey to the new structures might manifest as neutral effects in total fish abundance; and (iii) that fish stocking into populations that naturally reproduce in the lakes would not lead to additive effects on fish

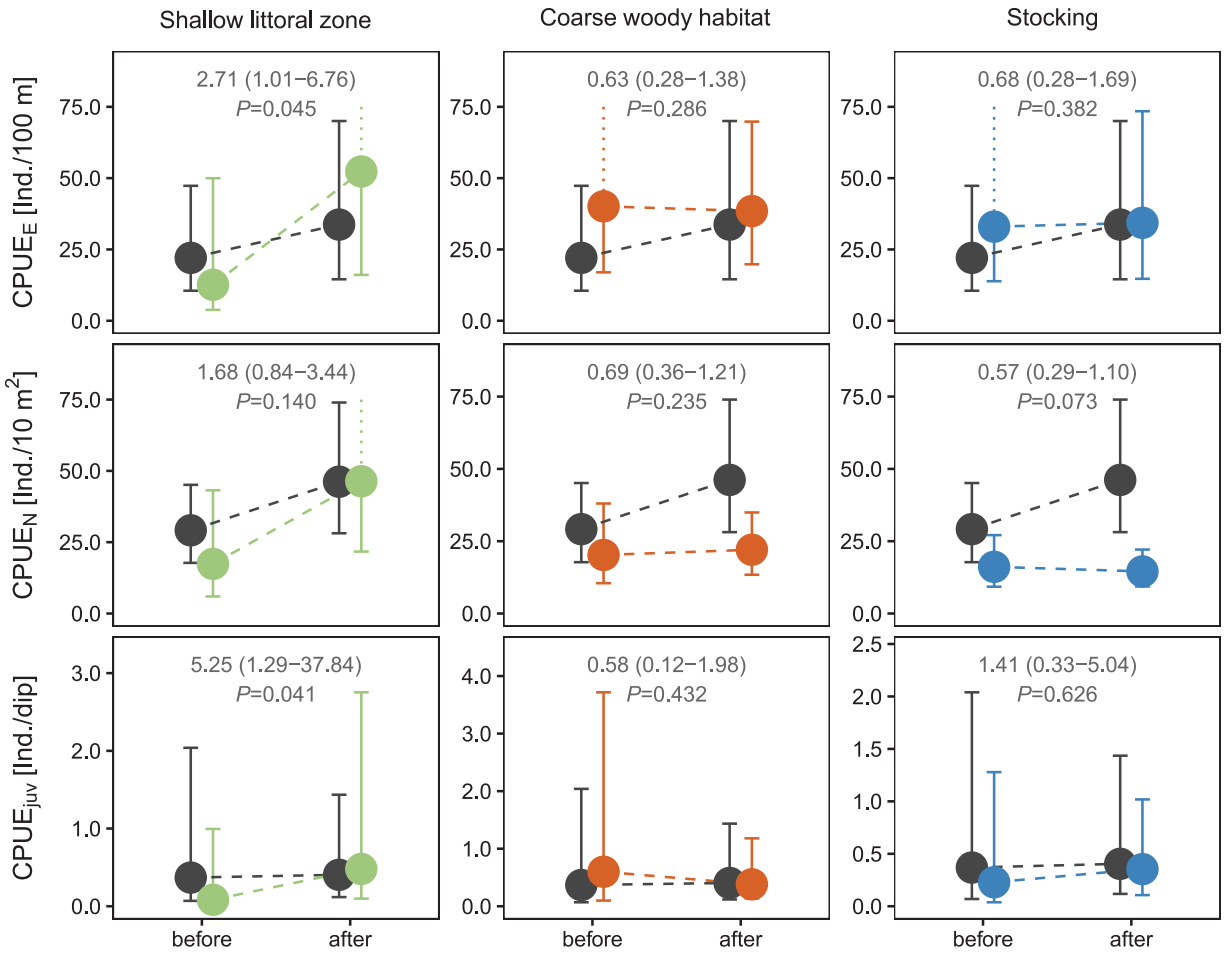


Fig. 2. Changes in total fish abundance in response to the three management measures and across three sampling methods. Colored circles indicate model-predicted mean CPUE before and after the management intervention. Dark circles indicate model-predicted CPUE for the control lakes. Error bars refer to the corresponding 95% bootstrapped CIs of the mean. Effect size estimates refer to the rate ratio of a given BACI interaction term (with 95% CI in parentheses).

abundance because of density-dependent mortality regulation.

Creating shallow littoral zones effectively enhances fish populations

Our study was based on a replicated and controlled set of experiments involving 20 gravel pit lakes (table S1 and fig. S1) and a sample of 159,943 fish captured 2 years before and 4 years after implementing three management interventions (*34*). Contrasting relative abundance (hereafter, abundance) changes between treated and unmanipulated control lakes (BACI design) (*34*) revealed that the creation of shallow littoral zones (12.5% increased littoral area on average; table S2 and figs. S2 and S3) was the most effective method to enhance fish populations (Fig. 2). Standardized total fish abundance (catch per unit effort, CPUE) as assessed by electrofishing ($CPUE_E$) increased significantly after shallow littoral zone creation by a factor of 2.71 compared with control lakes [generalized linear mixed-effects model (GLMM); confidence interval (CI) = 1.01 to 6.76, P = 0.045; effect controlled for coarse wood additions (*34*); table S3]. A similarly positive but nonsignificant trend was observed from gillnet-based abundance data ($CPUE_N$) in response to littoral zone creation and relative to controls (BACI effect = 1.68, CI = 0.84 to 3.44, P = 0.14; table S4). Juvenile fish (<100 mm) particularly benefited from the ecosystem-based management intervention of shallow littoral zone creation, showing a significant >5-fold abundance ($CPUE_{juv}$) increase compared with control lakes (BACI effect = 5.25, CI = 1.29 to 37.84, P = 0.041; table S5).

Our findings suggest that creating shallow littoral zones bolstered recruitment. The value of shallow littoral zones has long been recognized, particularly their importance during the life cycle of almost all temperate fishes (*21*, *35*, *36*). In view of the small spatial extent of littoral area enhancement (table S2), the clearly positive outcome of this measure is noteworthy. In particular, roach (*Rutilus rutilus*), one of the most frequent and abundant fish species of temperate European lakes, consistently increased in abundance in response to littoral zone creation, with significant effects detected for $CPUE_E$ and $CPUE_N$ (GLMM P < 0.05; Fig. 3 and tables S3 and S4), and a pronounced trend in $CPUE_{juv}$ (GLMM BACI effect = 5.00, P = 0.095; table S5). These positive effects associated with littoral zone creation were likely the result of enhanced reproduction and improved nursery function (*37*, *38*). However, ecosystem-wide benefits of littoral areas extend beyond providing suitable spawning grounds. In the presence of piscivores, littoral areas can provide beneficial foraging areas supporting juvenile growth through enhanced benthic production and warm water, and their vegetation cover and shallowness can effectively reduce vulnerability to predation (*18*, *21*, *39*). The profitable spatial overlap of forage and relatively predation-safe refuge areas likely contributed to enhanced juvenile fish development and the observed increase in total

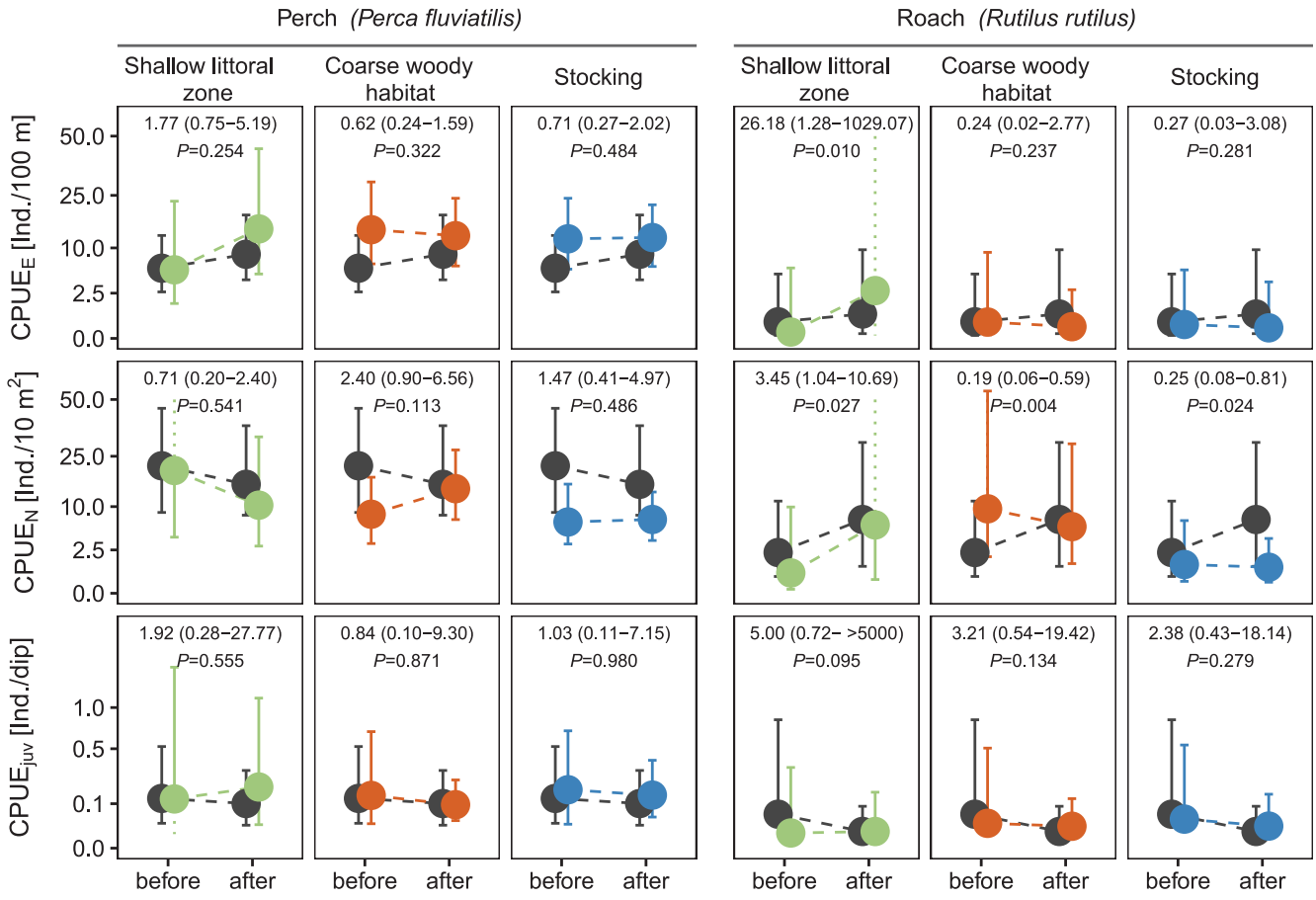


Fig. 3. Changes in abundance of perch and roach in response to the three management measures and across three sampling methods. Colored circles indicate model-predicted mean CPUE before and after the management intervention. Dark circles indicate model-predicted CPUE for the control lakes. Error bars refer to the corresponding 95% bootstrapped CIs of the mean. Effect size estimates refer to the rate ratio of a given BACI interaction term (with 95% CI in parentheses).

fish abundance after shallow littoral zone creation.

Convoluted effects of woody habitat additions

Although shallow zone creation was highly effective, habitat enhancement through coarse woody habitat additions to 21% of the shoreline (table S2 and fig. S3) alone did not significantly enhance total fish abundance or that of juveniles on average across all lakes (GLMM all $P > 0.05$; Fig. 2 and tables S3 to S5). Previous studies have shown that adding coarse wood to lakes is not necessarily associated with a short-term enhancement of fish populations (*23*, *24*). Responses to structural habitat enhancements in our study differed across sampling methods and lakes and between fish species, particularly between the two dominant fish species in gravel pit lakes, roach and European perch (*Perca fluviatilis;* hereafter, perch). In temperate lakes, both species form an important predator-prey relationship, with roach acting as an abundant zooplanktivorous fish and (large) perch as a key predator. In perch, coarse wood additions and the creation of littoral zones did not lead to significant abundance changes compared with the control lakes (GLMM all $P > 0.05$; Fig. 3 and tables S3 to S5). However, there was a strong but nonsignificant trend toward increasing perch $CPUE_N$ in response to coarse woody habitat additions (GLMM BACI effect = 2.40, CI = 0.90 to 6.56, $P = 0.113$; Fig. 3 and table S4). This trend suggests that wood additions could have increased perch mobility, as has been previously found in other piscivorous fish (*40*) [but see (*41*)], and thereby increased perch vulnerability to be caught by the passive sampling gear. Alternatively, coarse wood additions might have induced a spatial shift in perch from the littoral to the more open sublittoral caused by enhanced cover and increased foraging opportunities on prey at the edge of these structures (*30*, *40*). Spatial aggregation of fish near coarse woody habitats has repeatedly been observed (*23*, *42*) and was also suggested by our results showing a trend of relatively larger increases in perch abundance at sites that were closer to supplemented wood structures (GLMM BACI effect = 0.998, $P = 0.173$; table S6 and fig. S4).

In roach, coarse woody habitat additions resulted in a significant decrease of abundance in the sublittoral, as indicated by $CPUE_N$ compared with control lakes (BACI effect = 0.19, CI = 0.06 to 0.59, $P = 0.004$; Fig. 3 and table S4). However, there was no such evidence in roach $CPUE_E$ and $CPUE_{juv}$ (GLMM both $P > 0.05$; Fig. 3 and tables S3 and S5). The observed decline may have been caused by lower roach activity or the reduced use of sublittoral habitats after wood additions and in response to the elevated numbers in predators (*43*). Additionally, the refuge function of coarse woody habitats (*27*) might be less strong or universal and depends on its structural complexity (*28*, *43*). Supplemented woody habitats might have even facilitated predation, because predator-prey interactions and predation rates change with habitat structure and are often concentrated at the edge of complex habitats (*30*, *44*). In fact, structural habitats that simultaneously attract predator and prey might become ecological traps for the latter (*45*). Although fish might mistakenly perceive coarse wood as beneficial protective habitat, net-positive effects on

fish abundance may vanish because of increased predation rates. The attractiveness of structural habitats for fish that form ecological traps has previously been demonstrated, e.g., associated with artificial reefs (*45*).

Our study outcome might have been influenced by the uniformity of the supplemented wood bundles, which did not fully resemble the complexity of woody material originating from riparian trees (*25*, *42*), and by our rather short-term observation time frame. The first 4 years after manipulation might have been an insufficient amount of time to develop the full effects emerging from the long-lasting transformative processes of coarse wood, its colonization by invertebrates and periphyton, and its effects on stabilizing littoral habitats and macrophyte growth (*25*). Long-term processes might also explain why experimental coarse woody habitat removals often resulted in declines of single (prey) fish species (*37*, *46*) [but see (*47*)], whereas reciprocal whole-lake additions were frequently not associated with the short-term enhancement of fish populations [(*23*) and observed in this study: see Fig. 2].

Furthermore, the responses reported refer to average effects over multiple lakes that underwent a certain management measure compared with several control lakes, all characterized by inherent natural variability between ecosystems and in time. For example, the abundance of roach and perch increased in certain lakes after coarse wood additions, whereas it decreased in others (fig. S5), which resulted in a neutral mean effect over the sum of all studied lakes. Some have argued that whole-lake experiments using a low number of replicates of treatments with modest manipulative interventions but strong environmental noise might bear the risk of erroneously accepting the hypothesis of no treatment effect (*48*). This limitation cannot be fully excluded here despite the grand scale of our experiments with four (or more) replicate lakes manipulated at reasonable real-world levels. With our replicated study design, we aimed to account for lake- and year-specific effects of confounding factors (*34*) (e.g., differences in trophic state or specific weather events) that may have otherwise affected the results when solely relying on a single lake experiment. Nevertheless, or maybe for this very reason, we consider the positive effects emerging especially from ecosystem-based habitat management of lakes by shallow littoral zone creation as particularly robust and convincing.

No abundance-enhancing effects of fish stocking

Of the three management measures investigated, fish stocking achieved the poorest results. Stocking with five fish species (including prey and predator species) at 97 kg/ha did not produce any enhancing effects on fish abundance (Fig. 2), with total $CPUE_E$, $CPUE_N$, and $CPUE_{juv}$ remaining at similar levels after stocking (GLMM all $P > 0.05$; Fig. 2 and tables S3 to S5). When compared with controls, total $CPUE_N$ tended to even decrease after stocking (GLMM BACI effect = 0.57, CI = 0.29 to 1.10, $P = 0.073$; Fig. 2). At the species level, neither stocking of species reproducing [common bream (*Abramis brama*), tench (*Tinca tinca*), roach, and northern pike (*Esox lucius*)] nor of species not reproducing in the study lakes [pikeperch (*Sander lucioperca*)] resulted in increased fish abundance (GLMM all $P > 0.05$; table S7). Roach abundance $CPUE_N$ declined after stocking and when compared with unstocked control lakes (BACI effect = 0.25, CI = 0.08 to 0.81, $P = 0.024$; Fig. 3). Perch (not stocked) abundance did not change after stocking of other species and when compared with control lakes ($P > 0.05$; Fig. 3 and tables S3 to S5). These outcomes strongly agree with previous studies showing the very low additive effects of stocking (*12*–*14*). More specifically, for lakes that are at or bouncing around carrying capacity for juveniles, stocking additional fish simply increases competition for food and habitat without enhancing refuge opportunities from predation. Stocking may further lead to poorly conditioned fish or rapid recapture of stocked fish (*13*, *14*); replacement of wild with less fit, hatchery-produced fish (suffering from domestication selection) (*8*, *12*); and competitive disadvantages of stocked fishes, which might reduce the productivity of stocking-enhanced fish populations in the long term (*49*). We acknowledge that, as for other management measures, responses to fish stocking can be strongly context and species dependent and conditional on the specific lake characteristics and its species composition.

Conclusions

Our large-scale replicated study design that created 120 lake-years of observations revealed that ecosystem-based habitat management strongly outperforms the traditional, single-species–focused practice of fish stocking and single-structure–oriented management to support fisheries. Ecosystem-based habitat management, e.g., through the creation of shallow littoral zones (*20*, *34*), is the most promising management tool to enhance fish abundance in small lakes, especially when it effectively improves the growth-predation risk trade-off in early juvenile life stages. Holistic ecosystem-based management emphasizes the importance of improving both the essential ecosystem components and the societal dimension and decision-making processes (*1*). Our experiment was transdisciplinary and involved manipulations under community governance by local recreational fishing clubs and anglers, which likely contributed to a rethinking of stocking and fostered acceptance of more sustainable, ecosystem-based alternatives (*50*). In the UN Decade on Ecosystem Restoration, the implications of our work extend beyond fisheries toward conservation more generally. A focus on reestablishing central ecological processes and habitats is likely to have stronger long-term effects for the rebuilding of imperiled species than narrow, species-focused conservation actions.

REFERENCES AND NOTES

1. T. G. O'Higgins, M. Lago, T. H. DeWitt, *Ecosystem-Based Management, Ecosystem Services and Aquatic Biodiversity: Theory, Tools and Applications* (Springer, 2020).
2. J. Link, *Ecosystem-Based Fisheries Management: Confronting Tradeoffs* (Cambridge Univ. Press, 2010).
3. M. Ruckelshaus, T. Klinger, N. Knowlton, D. P. DeMaster, *Bioscience* **58**, 53–63 (2008).
4. C. J. Walters, in *Successes, Limitations, and Frontiers in Ecosystem Science*, M. L. Pace, P. M. Groffman, Eds. (Springer, 1998), pp. 272–286.
5. M. Palmer, A. Ruhi, *Science* **365**, eaaw2087 (2019).
6. S. R. Carpenter, in *Successes, Limitations, and Frontiers in Ecosystem Science*, M. L. Pace, P. M. Groffman, Eds. (Springer, 1998), pp. 287–312.
7. D. P. Armstrong, P. J. Seddon, *Trends Ecol. Evol.* **23**, 20–25 (2008).
8. K. Lorenzen, M. C. M. Beveridge, M. Mangel, *Biol. Rev. Camb. Philos. Soc.* **87**, 639–660 (2012).
9. G. G. Sass, A. L. Rypel, J. D. Stafford, *Fisheries* **42**, 197–209 (2017).
10. L. A. Eby, W. J. Roach, L. B. Crowder, J. A. Stanford, *Trends Ecol. Evol.* **21**, 576–584 (2006).
11. L. Laikre, M. K. Schwartz, R. S. Waples, N. Ryman; GeM Working Group, *Trends Ecol. Evol.* **25**, 520–529 (2010).
12. D. Hühn, K. Lübke, C. Skov, R. Arlinghaus, *Can. J. Fish. Aquat. Sci.* **71**, 1508–1519 (2014).
13. F. D. Johnston *et al.*, *Ecol. Appl.* **28**, 2033–2054 (2018).
14. K. Lorenzen, *Philos. Trans. R. Soc. London B Biol. Sci.* **360**, 171–189 (2005).
15. R. Arlinghaus, C. Riepe, S. Theis, T. Pagel, M. Fujitani, *J. Outdoor Recreat. Tour.* **38**, 100475 (2022).
16. R. Arlinghaus, K. Lorenzen, B. M. Johnson, S. J. Cooke, I. G. Cowx, in *Freshwater Fisheries Ecology*, J. F. Craig, Ed. (Wiley, 2016), pp. 557–579.
17. E. E. Werner, D. J. Hall, *Ecology* **69**, 1352–1366 (1988).
18. R. N. M. Ahrens, C. J. Walters, V. Christensen, *Fish Fish.* **13**, 41–59 (2012).
19. C. J. Walters, F. Juanes, *Can. J. Fish. Aquat. Sci.* **50**, 2058–2070 (1993).
20. J. L. Fischer, E. F. Roseman, C. M. Mayer, S. Qian, *Ecol. Eng.* **123**, 54–64 (2018).
21. I. J. Winfield, *Limnologica* **34**, 124–131 (2004).
22. R. G. Randall, C. K. Minns, V. W. Cairns, J. E. Moore, *Can. J. Fish. Aquat. Sci.* **53** (S1), 35–44 (1996).
23. G. G. Sass, S. R. Carpenter, J. W. Gaeta, J. F. Kitchell, T. D. Ahrenstorff, *Aquat. Sci.* **74**, 255–266 (2012).
24. K. E. Smokorowski, T. C. Pratt, *Environ. Rev.* **15**, 15–41 (2007).
25. M. Czarnecka, *Hydrobiologia* **767**, 13–25 (2016).
26. J. J. Taylor *et al.*, *Environ. Evid.* **8**, 19 (2019).
27. R. A. Everett, G. M. Ruiz, *Oecologia* **93**, 475–486 (1993).
28. J. P. Ziegler, C. J. Dassow, S. E. Jones, A. J. Ross, C. T. Solomon, *Can. J. Fish. Aquat. Sci.* **76**, 998–1005 (2019).
29. C. A. Layman, J. E. Allgeier, *J. Appl. Ecol.* **57**, 2139–2148 (2020).
30. G. G. Sass, C. M. Gille, J. T. Hinke, J. F. Kitchell, *Ecol. Freshwat. Fish* **15**, 301–308 (2006).
31. M. H. Olson *et al.*, *Fisheries* **23**, 6–12 (1998).
32. G. J. A. Hansen, J. W. Gaeta, J. F. Hansen, S. R. Carpenter, *Fisheries* **40**, 56–64 (2015).
33. A. P. Christie *et al.*, *J. Appl. Ecol.* **56**, 2742–2754 (2019).
34. Materials and methods are available as supplementary materials.
35. W.-C. Lewin, T. Mehner, D. Ritterbusch, U. Brämick, *Hydrobiologia* **724**, 293–306 (2014).
36. S. Matern, T. Klefoth, C. Wolter, R. Arlinghaus, *Hydrobiologia* **848**, 2449–2471 (2021).

37. J. W. Gaeta, G. G. Sass, S. R. Carpenter, . *Can. J. Fish. Aquat. Sci.* **71**, 315–325 (2014).
38. U. Kahl, S. Hülsmann, R. J. Radke, J. Benndorf, *Limnologica* **38**, 258–268 (2008).
39. V. Gotceitas, P. Colgan, *Oecologia* **80**, 158–166 (1989).
40. Q. C. Smith, G. G. Sass, T. R. Hrabik, S. L. Shaw, J. K. Raabe, *Ecol. Freshwat. Fish* **31**, 454–468 (2021).
41. T. D. Ahrenstorff, G. G. Sass, M. R. Helmus, *Hydrobiologia* **623**, 223–233 (2009).
42. M. G. Newbrey, M. A. Bozek, M. J. Jennings, J. E. Cook, *Can. J. Fish. Aquat. Sci.* **62**, 2110–2123 (2005).
43. C. S. DeBoom, D. H. Wahl, *Trans. Am. Fish. Soc.* **142**, 1602–1614 (2013).
44. J. F. Savino, R. A. Stein, *Environ. Biol. Fishes* **24**, 287–293 (1989).
45. S. E. Swearer *et al.*, *Front. Ecol. Environ.* **19**, 234–242 (2021).
46. G. G. Sass *et al.*, *Fisheries* **31**, 321–330 (2006).
47. K. E. Smokorowski, J. L. Pearce, W. D. Geiling, T. C. Pratt, *N. Am. J. Fish. Manage.* **41**, 142–157 (2021).
48. S. R. Carpenter, *Ecology* **70**, 453–463 (1989).
49. M. W. Chilcote, K. W. Goodson, M. R. Falcy, *Can. J. Fish. Aquat. Sci.* **68**, 511–522 (2011).
50. M. Fujitani, A. McFall, C. Randler, R. Arlinghaus, *Sci. Adv.* **3**, e1602516 (2017).
51. J. Radinger *et al.*, Data for: Ecosystem-based management outperforms species-focused stocking for enhancing fish populations, Figshare (2023); https://doi.org/10.6084/m9.figshare.20528295.

ACKNOWLEDGMENTS

We thank the Angler Association of Lower Saxony and all participating angling clubs for their participation; the chemical laboratory at IGB for analyses of water samples; the many helpful students, interns, and colleagues, especially A. Türck and R. Nikolaus, for their valuable assistance during field work; and S. R. Carpenter and two anonymous reviewers for their insightful feedback on the first draft of the manuscript. **Funding:** This study was jointly financed by the German Federal Ministry of Education and Research (BMBF) and the German Federal Agency for Nature Conservation (BfN) with funds granted by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU; grant no. 01LC1320A). **Author contributions:** R.A., T.K., and C.W. conceived and designed the study. S.M., F.F., and T.K. performed the research. J.R. and C.T.M. analyzed and visualized the data. J.R. and R.A. wrote the paper with substantial input from all coauthors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data and computer code used in the analysis are publicly available at the Figshare repository (*51*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf0895
Materials and Methods
Figs. S1 to S5
Tables S1 to S7
References (*52*–*70*)
MDAR Reproducibility Checklist

Submitted 13 October 2022; accepted 6 February 2023
10.1126/science.adf0895

By **Armando Andres Roca Suarez**

A story to remember

I walked into the conference room, eager to hear a presentation on the same topic as my ongoing Ph.D. project. But my enthusiasm didn't last long. As a scientist with attention-deficit/hyperactivity disorder (ADHD), the text-laden slides were difficult for me to process while I simultaneously tried to listen to what the presenter had to say. My mind began to wander, only to snap back to the talk when I heard the sound of clapping. Afterward, I was faced with awkward conversations during the coffee break as colleagues asked for my opinion on the research and I had no real answer to give them. The experience was frustrating, but it also helped me think about how I can craft my own presentations to interest all scientists—including audience members with ADHD.

I started taking notes during presentations, logging slide formats or delivery styles that made it easier for me to stay focused. Once I identified a potentially useful technique, I applied it during my own talks and asked for feedback from colleagues. Their positive comments reinforced to me that the techniques I found useful as an audience member helped and resonated with non-ADHD audience members as well. Here are some of those strategies.

"The techniques I found useful ... resonated with non-ADHD audience members as well."

LESS TEXT, MORE ILLUSTRATIONS. People living with ADHD are easily distracted and often find it challenging to keep their attention focused on a single task or stream of thought. Therefore, when presenting research, I try to make it as easy as possible for the audience to follow along with my slides as I speak. I keep them simple, using minimal text and inserting graphs, photos, and diagrams to illustrate my points as much as possible. Visual information takes less time to process, giving the audience an opportunity to be focused on what I am explaining.

TELL A COMPELLING STORY. Audience members will be more engaged if they're invested in finding out the answer to your scientific question, so the start and structure of a presentation are key. I try to craft introductions to highlight why the study is important and what gap in knowledge I'm trying to fill. If I have a surprising finding, I leave it for last. In this manner, the audience can follow a narrative arc that leads to a noteworthy conclusion. To emphasize the twists and turns in my story, I try to vary the tone and pace of my voice, using pauses to build tension. Nothing can make a person with ADHD lose attention faster than a monotone voice narrating a predictable story.

INTRODUCE FIGURES GRADUALLY. Although visuals are helpful, starting with a slide full of results gets my mind racing as it tries to process everything at the same time. To avoid that problem, I often start with relatively blank results slides, gradually adding more figures or diagrams in parallel with my explanations. Guide your audience through the variables and findings one step at a time.

UNDERSCORE YOUR MAIN MESSAGE. One of my favorite comedians, Chris Rock, once said during an interview that at the beginning of his career he worried the audience would forget the premise of the joke, so he started repeating it multiple times just to be sure. This developed into his signature delivery style. I've used that approach in my own presentations—spelling out the message of each slide in its title, repeating that message a second time in a bullet point at the bottom of each slide, and including a conclusions slide at the end of each talk that reiterates my main points. This ensures that even if someone's mind wandered for part of the talk, they'll still leave with the key takeaway message.

Most people in your audience won't have ADHD. But those who do will appreciate it if you use some or all of these tips. And I suspect others will as well. Humans love to hear compelling, easy-to-follow tales—so give your audience a story to remember. ■

Armando Andres Roca Suarez is a postdoctoral fellow at the Cancer Research Center of Lyon. Send your story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER